frontiers | Research Topics

Artificial intelligence in acute neurology

Edited by
Tarun Singh and Alejandro Rabinstein

Published in
Frontiers in Neurology

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-6342-7
DOI 10.3389/978-2-8325-6342-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Artificial intelligence in acute neurology

Topic editors

Tarun Singh — University of Michigan, United States

Alejandro Rabinstein — Mayo Clinic, United States

Citation

Singh, T., Rabinstein, A., eds. (2025). *Artificial intelligence in acute neurology*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6342-7

Table of contents

frontiers | Frontiers in Neurology

TYPE Editorial
PUBLISHED 29 April 2025
DOI 10.3389/fneur.2025.1605735

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Sean Ruland,
Loyola University Medical Center,
United States

*CORRESPONDENCE
Tarun D. Singh
✉ tarundsingh6@gmail.com

RECEIVED 03 April 2025
ACCEPTED 11 April 2025
PUBLISHED 29 April 2025

CITATION
Singh TD and Rabinstein AA (2025) Editorial: Artificial intelligence in acute neurology. *Front. Neurol.* 16:1605735. doi: 10.3389/fneur.2025.1605735

COPYRIGHT
© 2025 Singh and Rabinstein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Artificial intelligence in acute neurology

Tarun D. Singh[1]* and Alejandro A. Rabinstein[2]

[1]Department of Neurology and Neurosurgery, University of Michigan, Ann Arbor, MI, United States, [2]Department of Neurology, Mayo Clinic, Rochester, MN, United States

KEYWORDS

acute neurology, artificial intelligence, stroke, neurocritical care, neurohospitalist care

Editorial on the Research Topic
Artificial intelligence in acute neurology

Introduction

What if an algorithm could predict a life-threatening stroke hours before symptoms appear—or even before the patient realizes they are at risk? In acute Neurology, artificial intelligence (AI) is not just a tool; it can become the difference between life and death. No longer confined to theoretical algorithms or experimental models, AI is now influencing real-time decision-making, predictive analytics, and diagnostic accuracy. Its ability to analyze vast datasets, recognize complex patterns, and offer data-driven insights has transformed not just how we approach neurological emergencies, but also how we envision the future of patient care.

There is a lot of interest in AI's potential in automating image interpretation, identifying brain abnormalities with unprecedented accuracy, and accelerating diagnoses—thereby reducing the burden on clinicians and enhancing patient outcomes. This special edition, *"Artificial intelligence in acute neurology,"* reflects the rising global interest in this field and the growing momentum for integrating AI into clinical workflows. It features 17 high-quality manuscripts authored by 156 researchers from prestigious institutions across the United States, China, Germany, Japan, Singapore, India, the United Kingdom, Canada among other countries.

Since its launch, this Research Topic has garnered over 28,000 views and almost 10,000 downloads—demonstrating its global relevance and the increasing recognition of AI's transformative potential in clinical Neurology. Below, we highlight key insights, emphasize emerging trends, and discuss future directions for this dynamic and rapidly evolving field.

Key highlights from the research contributions

AI's applications in acute Neurology span a wide range—from prehospital care, predictive modeling and diagnostic innovations to clinical decision support. The manuscripts in this special edition collectively give a glimpse onto how AI can redefine neurological care across these domains.

Predictive models for neurological and systemic complications

AI-driven predictive models are enhancing clinicians' ability to anticipate complications such as early neurological deterioration, cerebral edema, and recurrent intracerebral hemorrhage. These tools enable earlier interventions, reducing morbidity, and improving long-term outcomes. Furthermore, AI's role extends beyond the brain—models predicting systemic issues like postoperative urinary retention and sepsis-associated encephalopathy illustrate AI's versatility in managing complex neurological patients holistically.

Diagnostic innovations using imaging data

AI has shown remarkable potential in imaging analysis, from detecting subtle ischemic changes on CT scans to predicting stroke outcomes using radiomics and machine learning. Models integrating clinical and imaging data have improved diagnostic precision for conditions like post-concussive syndrome and intracranial hemorrhage, even in resource-limited settings. These advancements are democratizing access to expert-level diagnostics globally.

AI in critical care and clinical decision-making

AI supports clinical decision-making by offering real-time, data-driven insights that enhance patient monitoring and risk assessment in Neurocritical care. Models predicting complications such as postoperative pneumonia in subarachnoid hemorrhage patients or identifying high-risk stroke patients underscore AI's potential to improve care pathways, reduce errors, and optimize resource allocation.

Prehospital and global applications

AI's influence extends beyond hospital walls. In prehospital settings, AI models may assist in rapid stroke classification and triage, improving time-to-treatment metrics. Notably, models designed for resource-limited environments highlight AI's potential to bridge healthcare disparities, enabling frontline providers to make informed decisions even in the absence of specialist support.

Collectively, these studies demonstrate AI's capacity to enhance neurological care across the entire continuum—from the prehospital environment to critical care units—ushering in an era where data-driven insights complement clinical expertise.

Ethical considerations in AI for acute neurology

While AI holds transformative potential, its integration into clinical practice raises critical ethical considerations. Addressing these challenges proactively is essential to ensure AI supports patient-centered, equitable, and safe neurological care.

Transparency and accountability

Many AI models operate as "black boxes," making it difficult to understand how specific predictions are generated. Ensuring transparency in AI algorithms is vital for building clinician trust. Explainable AI (XAI) techniques, such as SHapley Additive exPlanations (SHAP), help demystify these models by illustrating how various data inputs influence outcomes.

Bias and fairness

AI models are only as unbiased as the data they are trained on. If datasets lack diversity, models may perpetuate healthcare disparities, particularly for underrepresented populations. Ensuring fairness requires deliberate efforts to diversify training datasets and develop algorithms that perform consistently across different demographic groups.

Data privacy and security

AI relies on large datasets, often aggregated from multiple sources. Protecting patient privacy while enabling data-driven innovation is a delicate balance. Privacy-preserving techniques like federated learning offer promising solutions, allowing models to learn from decentralized data without compromising individual privacy.

Clinical responsibility

AI is a decision-support tool—not a replacement for clinical judgment. Clear guidelines must define the roles and responsibilities of clinicians when using AI-driven recommendations. In cases of adverse outcomes, questions around liability—whether it rests with the developer, the clinician, or the institution—must be addressed within legal and ethical frameworks.

By fostering transparency, fairness, and accountability, we can harness AI's potential responsibly, ensuring it complements rather than complicates clinical care.

Future directions for AI in acute neurology

As AI continues to evolve, several emerging trends are poised to shape the future of acute neurology.

Personalized neurological care

AI's ability to analyze large, complex datasets opens the door to truly personalized medicine. Integrating genomic, imaging, and clinical data, AI can predict individual responses to treatments, tailoring interventions for conditions like stroke, traumatic brain injury, and neuroinflammatory diseases.

Real-time decision support

The future lies in real-time AI applications that continuously monitor patients in neurocritical care units. These systems can detect early signs of clinical deterioration—such as impending cerebral edema, vasospasm in aSAH patients or intracranial hypertension in traumatic brain injury (TBI)—allowing for rapid interventions before irreversible damage occurs.

Adaptive learning models

Unlike traditional algorithms that remain static after deployment, adaptive AI models can learn from new data, continuously refining their predictions. This dynamic capability ensures that AI systems evolve alongside advances in medical knowledge and changes in patient populations.

Integration with brain-computer interfaces

The convergence of AI with BCIs offers exciting possibilities, particularly for patients with severe neurological impairments. AI-enhanced BCIs could improve communication and motor control for individuals with conditions like locked-in syndrome or advanced neurodegenerative diseases.

Global health and resource-limited settings

AI's potential to improve care extends beyond high-resource settings. Models designed for use in resource-limited environments—where access to specialists is scarce—can democratize neurological care, providing diagnostic and decision-support tools to frontline healthcare workers worldwide.

While the promise of AI in acute neurology is vast, realizing its full potential will require ongoing collaboration across disciplines, continuous evaluation, and a commitment to ethical, patient-centered care.

Call to action

The integration of artificial intelligence into acute neurology represents a paradigm shift—one that holds the promise of transforming patient outcomes through enhanced diagnostics, predictive analytics, and personalized care. However, realizing this potential requires coordinated efforts from clinicians, researchers, policymakers, and technologists.

Foster interdisciplinary collaboration

AI's successful integration into clinical practice depends on partnerships between neurologists, data scientists, ethicists, and engineers. Collaborative research will ensure that AI models are not only technically robust but also clinically meaningful.

Promote ethical AI development

Stakeholders must prioritize fairness, transparency, and accountability in AI development. Ethical frameworks should guide data collection, model training, and clinical deployment to ensure that AI benefits all patients equitably.

Invest in education and capacity building

Healthcare professionals need ongoing education to effectively leverage AI tools. Training programs should focus on understanding AI's capabilities and limitations, fostering a culture of informed adoption rather than passive reliance.

By embracing these principles, the neurological community can lead the way in shaping an AI-driven future that prioritizes patient wellbeing, clinical excellence, and ethical integrity.

Conclusion

The contributions in this special edition underscore the transformative potential of AI in acute neurology. From predictive modeling and diagnostic innovations to ethical considerations and future directions, the 17 manuscripts featured here reflect a vibrant, global effort to harness AI's power for the benefit of patients worldwide.

However, the journey toward full integration is not without challenges. Ethical dilemmas, data privacy concerns, and the need for interdisciplinary collaboration will shape how AI evolves within clinical practice. By addressing these challenges head-on, we can ensure that AI serves as a tool for empowerment—enhancing clinical decision-making, improving patient outcomes, and advancing the frontiers of neurological care.

We extend our deepest gratitude to the authors, reviewers, and editorial board members for their dedication and contributions. With over 16,000 views and 8,000 downloads, this Research Topic has already made a meaningful impact, sparking conversations, and inspiring innovation across the neurological community.

The future of acute Neurology will be defined not just by the data we collect and the AI models we develop, but primarily by how intelligently we use them. Now is the time to lead, innovate, and transform. And it is the time to decide how AI can work for us and not replacing us. We will always need empathetic and judicious neurologists. AI should support their clinical tasks and reduce cognitive biases, but for the

sake of our patients those tasks should remain an eminently human endeavor.

Author contributions

TS: Writing – review & editing, Writing – original draft. AR: Writing – review & editing, Writing – original draft.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 14 December 2023
DOI 10.3389/fneur.2023.1295642

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Giuseppe Miceli,
University of Palermo, Italy
Antonino Tuttolomondo,
Università degli Studi di Palermo, Italy

*CORRESPONDENCE
Nobukazu Miyamoto
✉ nobu-m@juntendo.ac.jp
Nobutaka Hattori
✉ nhattori@juntendo.ac.jp

RECEIVED 17 September 2023
ACCEPTED 01 December 2023
PUBLISHED 14 December 2023

CITATION
Miyamoto N, Ueno Y, Yamashiro K, Hira K, Kijima C, Kitora N, Iwao Y, Okuda K, Mishima S, Takahashi D, Ono K, Asari M, Miyazaki K and Hattori N (2023) Stroke classification and treatment support system artificial intelligence for usefulness of stroke diagnosis.
Front. Neurol. 14:1295642.
doi: 10.3389/fneur.2023.1295642

COPYRIGHT
© 2023 Miyamoto, Ueno, Yamashiro, Hira, Kijima, Kitora, Iwao, Okuda, Mishima, Takahashi, Ono, Asari, Miyazaki and Hattori. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Stroke classification and treatment support system artificial intelligence for usefulness of stroke diagnosis

Nobukazu Miyamoto[1]*, Yuji Ueno[1], Kazuo Yamashiro[1], Kenichiro Hira[1], Chikage Kijima[1], Naoki Kitora[2], Yoshihiko Iwao[2], Kayo Okuda[2], Shohei Mishima[2], Daisuke Takahashi[2], Kazuto Ono[3], Mika Asari[4], Kazuki Miyazaki[4] and Nobutaka Hattori[1]*

[1]Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan, [2]HACARUS INC., Kyoto, Japan, [3]Ohara Pharmaceutical Co., Ltd., Tokyo, Japan, [4]PARKINSON Laboratories Co., Ltd., Tokyo, Japan

Background and aims: It is important to diagnose cerebral infarction at an early stage and select an appropriate treatment method. The number of stroke-trained physicians is unevenly distributed; thus, a shortage of specialists is a major problem in some regions. In this retrospective design study, we tested whether an artificial intelligence (AI) we built using computer-aided detection/diagnosis may help medical physicians to classify stroke for the appropriate treatment.

Methods: To build the Stroke Classification and Treatment Support System AI, the clinical data of 231 hospitalized patients with ischemic stroke from January 2016 to December 2017 were used for training the AI. To verify the diagnostic accuracy, 151 patients who were admitted for stroke between January 2018 and December 2018 were also enrolled.

Results: By utilizing multimodal data, such as DWI and ADC map images, as well as patient examination data, we were able to construct an AI that can explain the analysis results with a small amount of training data. Furthermore, the AI was able to classify with high accuracy (Cohort 1, evaluation data 88.7%; Cohort 2, validation data 86.1%).

Conclusion: In recent years, the treatment options for cerebral infarction have increased in number and complexity, making it even more important to provide appropriate treatment according to the initial diagnosis. This system could be used for initial treatment to automatically diagnose and classify strokes in hospitals where stroke-trained physicians are not available and improve the prognosis of cerebral infarction.

KEYWORDS

stroke, TOAST classification, multimodal artificial intelligence, k-Nearest Neighbor method, leave-one-out cross-validation method, cerebral infarction

Highlight

- SCTSS-AI classifies strokes with over 85% accuracy, aiding in treatment decisions.
- Utilizing multimodal data, the AI provides explanations and improves prognosis of cerebral infarction.
- The system addresses the shortage of stroke experts, enabling automatic diagnosis and classification.

Introduction

Cerebrovascular disease, commonly referred to as stroke, is a leading cause of death and chronic disability on a global scale (1–3). Approximately 80% of strokes are caused by cerebral ischemia (4). In addition, energy depletion and cell death can cause ischemic brain injury (5). These injuries lead to functional impairment of the injured neurons, leading to severe long-term disability. In the initial diagnosis, brain imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI), to detect tissue necrotic areas of cerebral infarction are important tools for ischemic stroke assessment (6).

Treatment of ischemic stroke includes intravenous thrombolysis, intra-arterial therapy, and mechanical revascularization (7). Although it is possible to diagnose stroke without being a stroke specialist, stroke-trained physicians classify stroke severity using knowledge of how the physiology of different stroke types is reflected in image textures (8). However, manual image analysis is labor intensive (8) and prone to inter- and intra-operator variability (9, 10). Furthermore, expert analysis is limited by the number and areas where specialists practice (11), which results in increased diagnostic costs. Automatic lesion identification and subsequent stroke severity classification can significantly reduce drawing time and accurately detect lesions (11). The development of computer-aided detection and diagnostic systems based on the automatic detection of post-stroke brain lesions is an active research field. In such studies, research is being conducted to construct an automated stroke severity classification system using either CT or MRI. Both methods yield a graphical representation of the human brain containing distinct image objects. Identifying such objects through image segmentation is an important step in extracting diagnostically important information. CT is faster and less expensive and more widely used globally than MRI. However, MRI is suitable for constructing an automated stroke severity classification system because MRI is much more sensitive for acute ischemic lesions than CT (12) and MRI scans can be enhanced by adding functional information to the anatomical data to form diffusion-weighted images (DWI).

There are regional disparities in the number of physicians who can diagnose stroke accurately globally. In Japan, there are many stroke-trained physicians in urban areas; however, there are fewer in rural hospitals. Thus, initial stroke treatment is provided by general physicians who are not trained in stroke care (13). To solve this problem, the Japan Stroke Association has provided guidelines on "drip-and-ship treatment," but this only increases the burden on urban stroke-trained physicians (14). In stroke treatment, it is important to classify the acute phase of stroke and treat patients according to the stroke classification, even in environments where mechanical thrombectomy and intravenous thrombolysis are not available (15). However, differences in functional prognosis have been reported between patients treated by stroke specialists and those treated by general physicians (16). We aimed to develop an artificial intelligence (AI)-based stroke diagnosis aid system using MRI to automatically diagnose and classify strokes in hospitals where stroke-trained physicians are not available, and to link this to initial medical care.

Patients and methods

Patients

We developed a Stroke Classification and Treatment Support System AI (SCTSS-AI) equipped with the infarct detection AI and the stroke classification AI for cerebral infarction. The development was approved by the Human Ethics Review Committee of Juntendo University School of Medicine (E22-0028). The stroke classification AI was established using the medical records and MRI data of Cohort 1, who were admitted to Juntendo University Hospital's Neurology Department between January 2016 and December 2017 for cerebral infarction or developed cerebral infarction while admitted and were treated at the Neurology Department (Figure 1A). The infarct detection AI was trained primarily using MRI data from patients in Cohort 1 with the three main types of Trial of Org 10,172 in Acute Stroke Treatment (TOAST) classification. To confirm the accuracy of SCTSS-AI, we used another data set provided from Cohort 2, who were treated at the same institution as Cohort 1 for stroke between January 2018 and December 2018 (Figure 1A).

The exclusion criteria for both cohorts were as follows: 1) patients aged <20 years at stroke onset, 2) patients with stroke >8 days after stroke onset, 3) patients who had not undergone MRI, 4) patients with stroke of undetermined etiology (negative evaluation and two or more causes identified), 5) cases diagnosed with aortic arterial dissection, and 6) patients who were judged by three stroke experts to be ineligible for data analysis. The diagnostic results and treatment methods provided by the system were constructed in accordance with the Japan Stroke Treatment Guidelines 2021 (14) and in the final evaluation, training data and test data were completely separated to evaluate the generalization performance of the system.

Collected data set

We extracted the following information from the medical records of each patient to establish SCTSS-AI: 1) demographic data; 2) vital signs at presentation and laboratory findings including ECG, fibrin/fibrinogen degradation products [FDP] D dimer, brain natriuretic peptide [BNP], N-terminal pro-BNP [NT-proBNP], estimated glomerular filtration rate, and high-sensitivity C-reactive protein on admission; 3) medications taken upon admission, with particular attention paid to anti-platelets, anti-coagulants, anti-hypertensives, and statins; 4) vascular risk factors for stroke, such as hypertension (HT; systolic blood pressure [BP] > 140 mmHg, diastolic BP > 90 mmHg, or drug treatment for HT), dyslipidemia (DL; defined as low-density lipoprotein [LDL] cholesterol level of >140 mg/dL, high-density lipoprotein [HDL]-cholesterol level of <40 mg/dL, triglyceride [TG] level of >149 mg/dL, or drug treatment for DL), diabetes mellitus (DM; defined as glycated hemoglobin level of >6.4%, or drug treatment for DM), a cardioembolic source according to TOAST classification (17), transient ischemic attack, and smoking history (as reported by the patient and their family); 5) stroke mechanism according to TOAST criteria (17); and 6) baseline National Institutes of Health Stroke Scale (NIHSS) score (18), as recorded by stroke-trained neurologists that were certified in the application of the NIHSS, on admission. Brain CT/MRI and

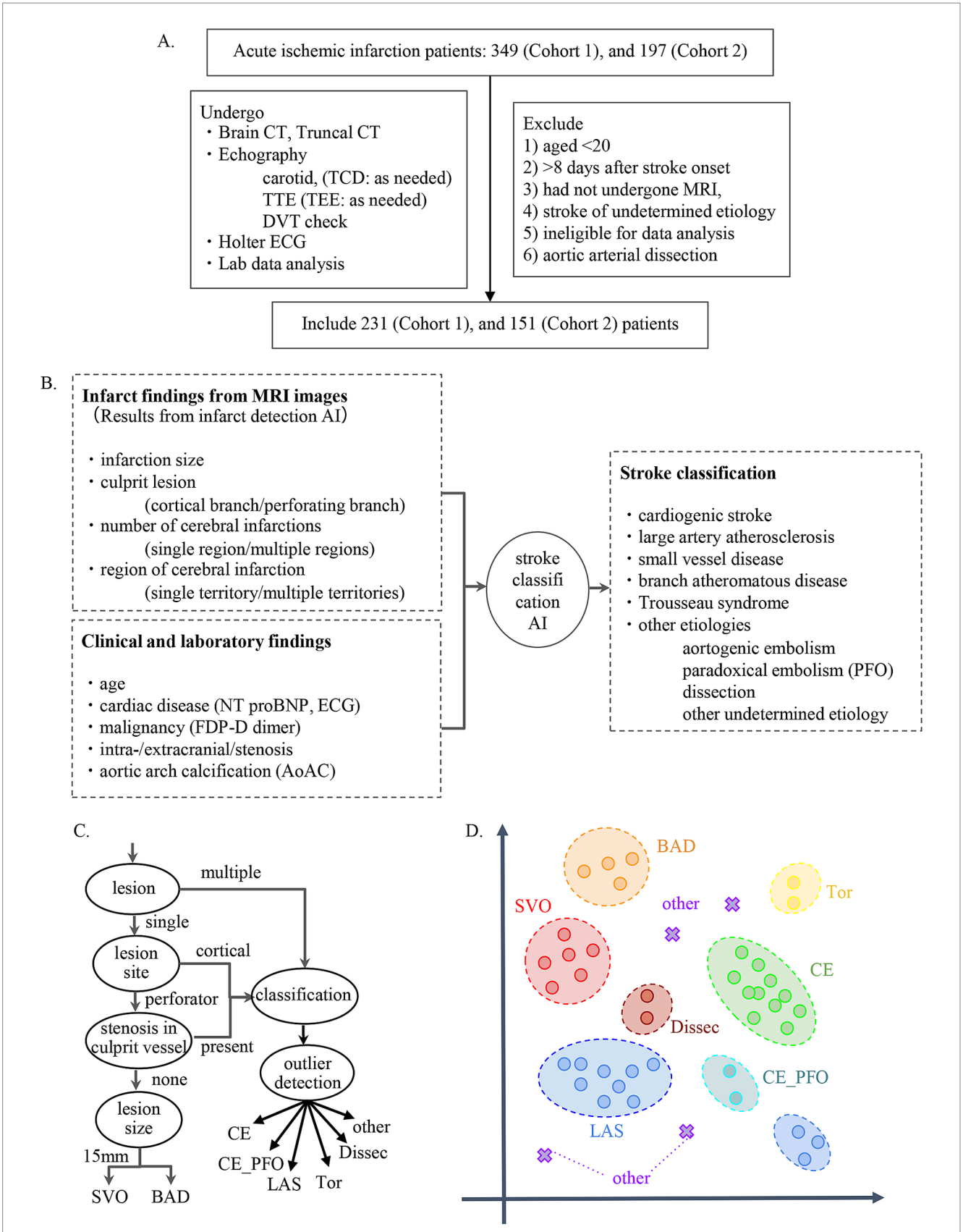


FIGURE 1
(A) Flow chart describing enrollment of patients with stroke in the present study. **(B)** Flowchart for establishment of stroke classification AI. **(C)** Decision flow for developing stroke classification AI. **(D)** k-Nearest Neighbor method. The k-Nearest Neighbor classification algorithm was used to classify SVO, BAD, LAS, and CE. For the other determined etiology, a method combining the concept of abnormality detection was used.

electrocardiography were performed in all patients, and we diagnosed brain infarction by focal hyper-intensity that was judged not attributable to normal anisotropic diffusion or magnetic susceptibility artifact.

Establishment of infarct detection AI

Adjustment of MRI images

We constructed the infarct detection AI that was used to derive the features that determine the stroke classification using MRI from a variety of different resolutions and manufacturers to assess patients who were initially suspected of stroke between 6 h and 7 days from the onset time as training data. The process of correcting the signal values of the MRI images based on the positions of the peaks was performed. To correct for differences in image orientation and position, we used image processing to measure the orientation of the head and perform rotation correction to adjust the tilt and to correct for differences in imaging range in the slice direction (Z-axis) and brain size, and used the Dynamic Time Warping technique to correct the Z-axis position (19). Variations in images among cases were corrected.

Identification of features

Using the corrected image data, the infarct detection AI derives the following features and provides them to SCTSS-AI. The stroke classification AI is designed to diagnose cerebral infarction based on TOAST classification and to propose treatment methods (Figure 1B). Infarct-related features were as follows: (i) size (< 1.5 cm/1.5 cm or larger), (ii) culprit lesion (cortical branch/perforating branch as the preferred site of small vessel occlusion [SVO]/branch atheromatous disease [BAD]), (iii) number of cerebral infarctions (single region/multiple regions), (iv) region of cerebral infarction (single territory/multiple territories), (v) presence of intracranial stenosis, (vi) presence of carotid artery stenosis, and (vii) presence or absence of cardiac disease as a risk factor for embolic source (including NT-proBNP). Of these, (i)-(iv) were extracted using infarct detection AI. The infarct detection AI was built using DWI and apparent diffusion coefficient (ADC) mapping, which are used to distinguish between infarcts and artifacts. The design of the features to be used as machine learning input was based on domain knowledge about the difference between infarcts and artifacts. The 3D positional and symmetry information of candidate pixels were used as features in the construction of the infarct detection AI, based on the artifacts tending to occur at specific locations and symmetrically. To distinguish between the feature of high signal at the infarction point of DWI, and cases where the high signal is not due to the infarction point but to the effect of T2 shine through, the pixel value of the ADC map and the amount of information of surrounding pixel values were also used as feature values (Figure 2A). Designing input features that leverage domain knowledge is difficult to incorporate into deep learning-based machine learning and is one of the major differences between the AI that we built and deep learning-based AI.

Annotation MRI images

To build the infarct detection AI to derive the above features, stroke experts with more than 10 years of experience annotated the MRI dataset as training data.

Construction of stroke classification AI

SCTSS-AI was designed so that the infarct detection AI and the stroke classification AI work in tandem to classify stroke. The infarct detection AI extracts the features from the MRI data, and the stroke classification AI combines infarct-related features and stroke classification-related features from the medical records to make a diagnosis.

Selection of features

The stroke classification AI was constructed using the patients' background, medical history, and clinical and laboratory findings used to classify stroke. The following additional features related to stroke classification were selected by 231 of the cases provided by Juntendo Hospital (Figure 1A): (viii) malignancy (treated/not treated), (ix) D-dimer, (x) grade of aortic arch calcification (AoAC) by chest X-ray (grade 0–3), and (xi) age. Features were selected based on TOAST criteria and our domain knowledge. These features were reported in previous studies and also confirmed in our analysis. It was noted that the blood fibrin degradation products, such as D-dimer, tend to be higher in Trousseau syndrome than in other stroke types due to hypercoagulability caused by malignancy, that AoAC tends to be higher in aortic primary cerebral embolism than in other stroke types (20), and that cardiogenic cerebral embolism due to patent foramen ovale (PFO) and arterial dissection are common stroke types in young patients (Figure 3).

Algorithm design

We designed an algorithm that makes inferences if part of the information is missing, as some tests necessary for diagnosis are not performed immediately after MRI in the clinical setting. The AI executes an appropriate model corresponding to features with missing values. The algorithm for diagnosis of cerebral infarction was constructed based on the TOAST classification approach (Figure 1C). This algorithm was based on the long-used algorithm of Lee et al. (21), which was modified to incorporate TOAST classification from imaging. In addition, we employed the k-Nearest Neighbor (kNN) method to classify stroke other than Others, and combined kNN with anomaly detection for Others (Figure 1D).

Evaluation methods & statistical analysis methods

The performance of the AI constructed from the 231 cases of Cohort 1 was assessed by the Leave One Out Cross Validation (LOOCV) method (22). Out of the 231 cases in Cohort 1, one case was extracted as the data for evaluation, and stroke was classified using the stroke classification AI trained with the remaining data and compared with the stroke classification determined by the stroke experts. This evaluation was repeated until all 231 cases were used as data for evaluation, and the stroke classification AI was evaluated. Final accuracy of the stroke classification AI with verified performance using LOOCV was evaluated using 151 independent cases from Cohort 1 (Cohort 2).

All data analysis, including image processing and feature extraction of the infarct detection AI, training and evaluation of the stroke classification AI, and visualization were performed under the Python 3.7 environment. The performance metrics used in the evaluation of the stroke classification AI were: accuracy sensitivity, precision, and *F* value expressed by the following formulas:

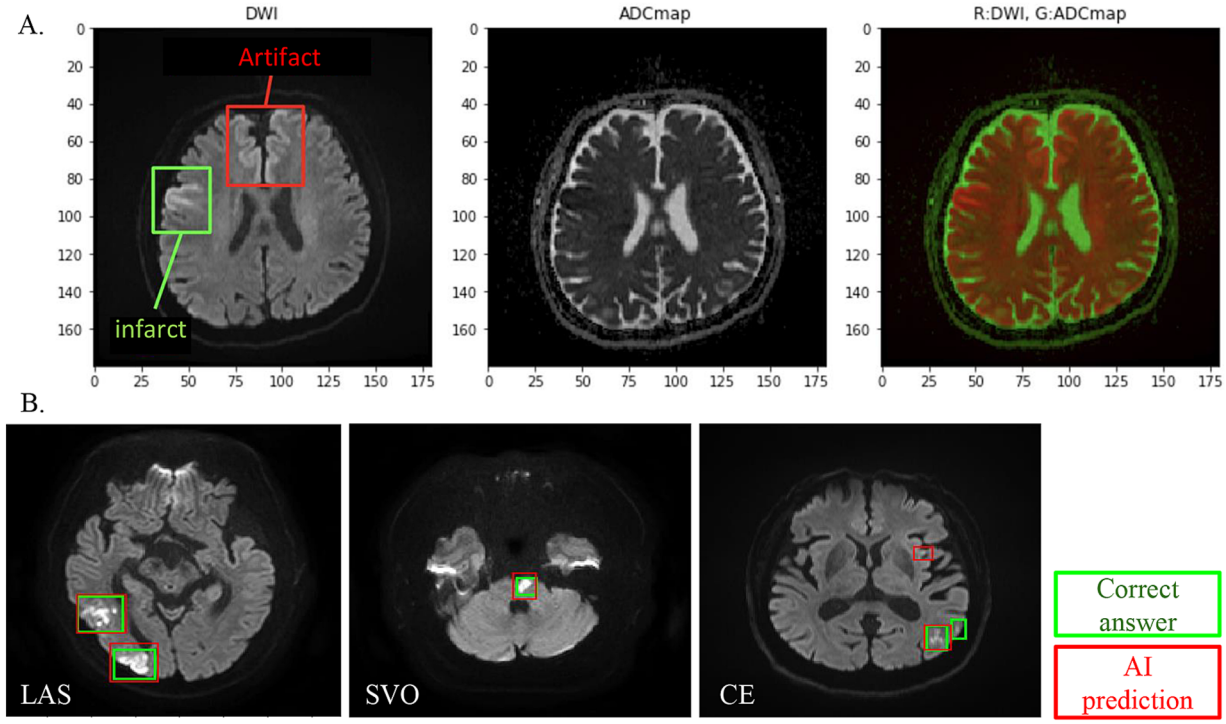


FIGURE 2
(A) DWI and ADC combined analysis method for infarct detection AI visualizes and evaluates the difference between DWI high-intensity area and ADC low-intensity area, except for symmetrical high-signal areas and areas where artifacts are likely to appear. **(B)** Infarct detection AI overview. One hundred cases were studied, and 18 cases were evaluated. In infarct lesion units, the sensitivity (recall) was 80% and the match rate (lesion-by-lesion evaluation) was 84%.

$$TP + FN = Patient\ number_{stype}$$

$$TP + FP = Predicted\ number\ of\ cases_{stype}$$

$$TP = Number\ of\ correct\ cases_{stype}$$

$$sensitivity_{stype} = \frac{Number\ of\ correct\ cases_{stype}}{Patient\ number_{stype}},$$

$$precision_{stype} = \frac{Number\ of\ correct\ cases_{stype}}{Predicted\ number\ of\ cases_{stype}}$$

$$F\,value_{stype} = \frac{2^{*}\,sensitivity_{stype}{}^{*}\,precision_{stype}}{sensitivity_{stype} + precision_{stype}},$$

$$accuracy = \frac{\sum_{stype} Number\ of\ correct\ cases_{stype}}{\sum_{stype} Patient\ number_{stype}}$$

where TP is True Positives, FN is False Negatives, FP is False Positives, and stype is subtype. Sensitivity, precision, and *F* value were calculated for each subtype. Accuracy was calculated for all patients. Sensitivity indicates how many patients with a subtype were actually detected as patients. Precision indicates how correct the predicted result was. F value is an integrated value of sensitivity and Precision and can be evaluated considering their trade-off. Accuracy indicates how many patients overall were predicted as having the correct disease type.

Statistical analysis

The data were analyzed with SPSS 29.0 (SAS Institute Inc., Cary, NC). Data are expressed as mean ± standard deviation values for continuous variables. All statistical analyzes were performed using χ^2 test for categorical variables, t-test for parametric analyzes. *p*-values of < 0.05 were considered significant.

Results

As cohort 1, 231 from 278 people, 151 from 197 people as cohort 2 were enrolled. Background factors and examination data of cohort 1 and 2 patients revealed no difference between cohort 1 and 2, except diastolic blood pressure, heart rate, stroke classification, blood sugar, HbA1c, triglyceride, eGFR on arrival (Table 1).

AI for infarct detection

We evaluated the infarct detection AI that determines the presence of absence of infarction using a machine learning algorithm

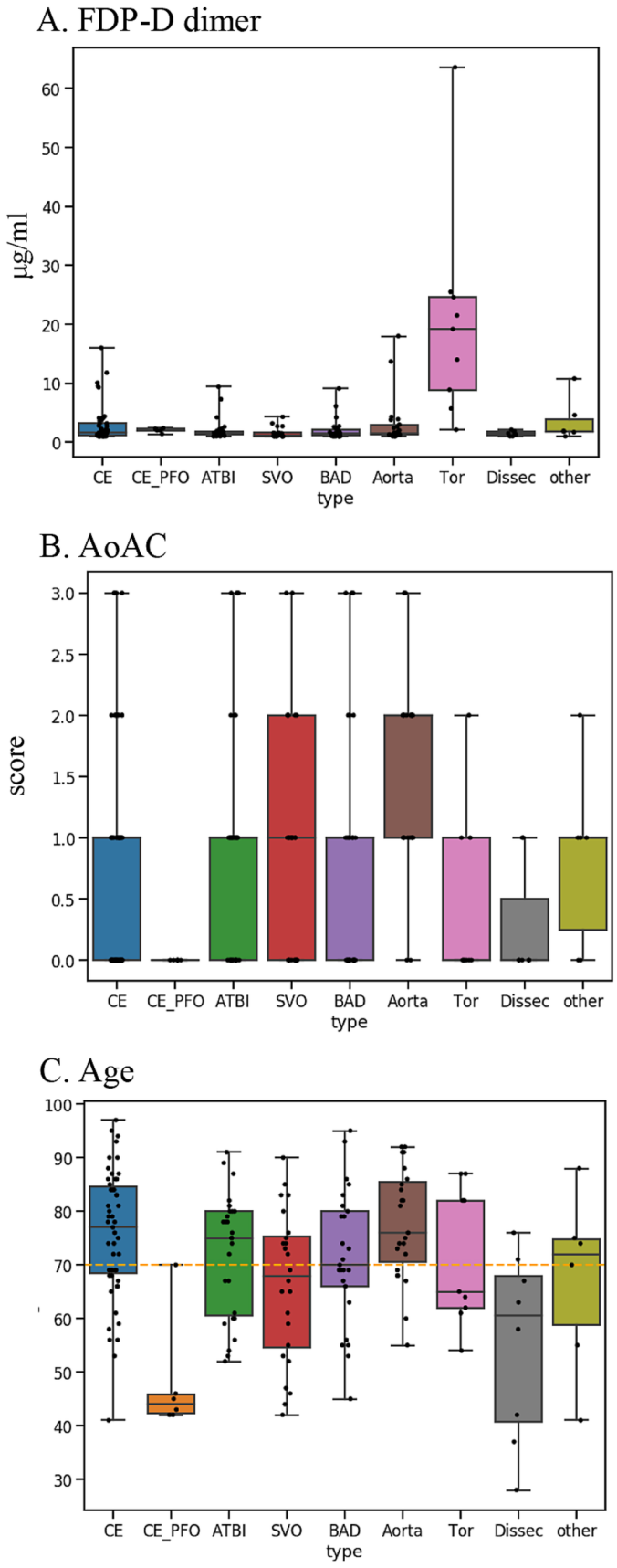


FIGURE 3
Key features used for disease typing AI: **(A)** D-dimer. **(B)** AoAC [Shimada et al. (17)]. **(C)** Age. AoAC increased in aortogenic embolism, and patients with CE-PFO/dissection were younger than other etiologies. FDP-D dimer was higher in patients with Trousseau syndrome.

10.3389/fneur.2023.1295642

TABLE 1 Background factors, stroke classification and examination data.

	Cohort 1 (231)		Cohort 2 (151)		*p*-value
	N	%	*N*	%	
Sex (male)	154	66.7	94	62.2	0.377
Age	69.4 ± 14.8		70.9 ± 14.2		0.172
Body height (cm)	161.4 ± 9.8		160.4 ± 9.3		0.250
Body weight (kg)	60.8 ± 13.3		59.0 ± 14.0		0.107
BMI (kg/m^2)	23.2 ± 3.8		22.6 ± 3.72		**0.047**
Systolic blood pressure on arrival (mmHg)	150.3 ± 29.8		153.3 ± 28.2		0.220
Diastolic blood pressure on arrival (mmHg)	82.2 ± 17.7		85.4 ± 18.0		**0.049**
Heart rate on arrival (/min)	76.1 ± 15.4		79.8 ± 14.9		**0.050**
Smoking habit	64	27.7	29	19.2	0.058
HT	150	64.9	111	73.5	0.078
DM	53	22.9	58	38.4	**0.001**
DL	79	34.1	101	66.8	**<0.001**
Ischemic heart disease	22	9.5	14	9.2	0.934
Af	39	16.8	37	24.5	0.068
Active Malignancy	18	7.79	13	8.6	0.775
Stroke classification					**0.005**
SVO	25	10.8	19	12.5	
BAD	28	12.1	14	9.2	
LAS	29	12.5	25	16.5	
CE	51	22.0	52	34.4	
CE-PFO	7	3.0	1	0.6	
aortogenic embolism	40	17.3	18	11.9	
Trousseau syndrome	9	3.8	9	5.9	
Dissection	15	6.4	8	5.2	
other	27	11.6	5	3.3	
Laboratory data on arrival					
WBC (/μL)	7,596 ± 3,490		7,194 ± 2,919		0.126
PT-INR	1.08 ± 0.22		1.10 ± 0.19		0.305
FDP-D dimer (μg/mL)	3.24 ± 7.08		4.15 ± 7.29		0.118
Blood sugar (mg/dL)	124.5 ± 49.2		138 ± 65.3		**0.009**
HbA1c (%)	6.15 ± 1.05		6.57 ± 1.71		**0.002**
LDL (mg/dL)	116.4 ± 36.8		116.5 ± 49.3		0.496
HDL (mg/dL)	51.3 ± 15.0		50.3 ± 16.5		0.270
TG (mg/dL)	120.6 ± 72.6		143.5 ± 104.2		**0.009**
UA (mg/dL)	5.50 ± 1.44		5.48 ± 1.70		0.439
Cre (mg/dL)	0.945 ± 0.984		1.063 ± 1.144		0.149
eGFR (mL/min/1.73 m^2)	75.53 ± 30.72		65.02 ± 26.13		**<0.001**
hsCRP (mg/dL)	1.262 ± 3.277		1.598 ± 3.755		0.186
NT-proBNP (pg/dL)	1051.88 ± 2658.45		1302.85 ± 4150.58		0.266

Bolded numbers indicate $p < 0.05$ compare to cohort1 and cohort2. CE, cardiogenic embolism; CE-PFO, cerebral embolization through a patent foramen ovale (paradoxical embolism); LAS, large artery atherosclerosis; SVO, small vessel occlusion; BAD, branch atheromatous disease.

called gradient boosting decision tree using 18 cases in Cohort 1 (SVO: 4 cases; cardiogenic embolism [CE]: 10 cases; large artery atherosclerosis [LAS]: 4). We found sensitivity (reproducibility: the percentage of infarcts that the AI could detect as infarcts out of those that were actually infarcts) of 80 and 84% (lesion-by-lesion evaluation). Thus, our infarct detection AI had a sensitivity (recall) of 80%, with few undetected infarcts and few false detections of infarcts (Figure 2B).

AI for stroke classification

The results of the evaluation of the stroke classification AI trained using features of 231 cases: SVO/BAD: 53 cases, CE: 51 cases, LAS: 29 cases, Trousseau syndrome: 9 cases, and Others, including aortogenic embolism, paradoxical embolism, and other undetermined etiology (embolic stroke of undetermined sources and cerebral artery dissection [ESUS+D]: 89 cases) in Cohort 1 using the LOOCV method showed an 88.7% correct rate (Tables 2, 3).

In the present evaluation, high accuracy was obtained for each stroke classification, especially for CE, LAS, and BAD/SVO; however, when attempting to predict BAD/SVO and ESUS+D more finely, the accuracy rate was lower (correct rate of 76.6%, Tables 4, 5). Because SVO/BAD is diagnosed within 1–3 days of the onset of the disease, it is difficult to distinguish SVO/BAD using the information available in the emergency room. Furthermore, it is impossible to diagnose paradoxical embolism unless transesophageal echocardiography is performed. On the other hand, if the patient has been treated for carcinoma, and the FDP D-dimer is elevated (Figure 3A), Trousseau syndrome could be allowed for differentiation. Another problem is that accurate reading of magnetic resonance angiography (MRA) is often difficult for non-stroke physicians. In the acute phase of treatment, the classification into CE, LAS, BAD/SVO, Trousseau syndrome, and ESUS+D may be more practical.

Verification of stroke classification AI

We evaluated the diagnostic accuracy of the stroke classification AI constructed using Cohort 1 using Cohort 2 (total 151 cases). The correct rate for stroke classification (SVO/BAD: 33 cases, CE: 52 cases, LAS: 25 cases, Trousseau syndrome: 9 cases, and ESUS+D: 32 cases) was 86.1% (Tables 6, 7), which was similar to the accuracy rate of Cohort 1 using the LOOCV method.

Discussion

In this study, we established SCTSS-AI for cerebral infarction diagnosis that uses image-based infarction detection and medical data to determine the stroke classification. We found that machine learning with the incorporation of clinical information can diagnose cerebral infarction based on the TOAST classification, which could not have been achieved with image-based stroke diagnosis AI alone. The high performance of this SCTSS-AI despite the diversity of MRI images used for construction, and the fact that there is no current AI that can classify the stroke types to provide appropriate treatment according to the initial treatment of cerebral infarction, suggest that this AI can be used in actual clinical practice in the future. SCTSS-AI can make inferences even when some features of data have missing values for a stroke classification, provided that in this verification work, the model was analyzed using cases with no missing data from a single institution. Further studies are needed to investigate whether the model can be analyzed even with missing data, and whether it can be used with high accuracy even when data from multiple institutions are used.

There are different stroke types that require different clinical management. Therefore, classification of stroke types is necessary for early treatment and prevention (23). Subudhi et al. (24), evaluated DWI with a support vector machine classifier according to the Oxfordshire Community Stroke Project (OCSP) classification. They obtained an accuracy of 92.9%, sensitivity of 90.4%, and specificity of 93.3% in differentiating among total anterior circulation infarction, partial anterior circulation infarction, and lacunar circulation infarction. However, posterior circulation infarction, another subclass of OCSP, has not been evaluated, and the MobileNetV2 convolutional neural network (CNN) model, which was fine-tuned to classify cerebral infarcts according to vascular territory, had an accuracy of 93% (25). However, only subtypes covering 75–80% have been evaluated, and this analysis requires the use of a CNN model, which requires a large number of patient cases. The solution to this problem was the creation of the ImageNet dataset, with over 15 million images labeled in 22,000 different categories (26). ImageNet is often used to measure the accuracy of current CNN models. EfficientNet and MobileNetV2 CNN models were preferred for transfer training compared to similar models because of their lower computational load and ImageNet's higher accuracy (27, 28). However, the advantage of our method is that by inputting MRI, laboratory data, X-ray data, ECG data, and other data used in daily medical treatment of patients into the application, stroke classification can be accurately made, and appropriate treatment can be immediately initiated for patients.

Current AI advancements in stroke TOAST classification are focused on predicting prognosis after the onset of stroke (29). The LAS diagnostic criterion requires over 50% stenosis in proximal arteries (30). A previous study reported the use of computer-based diagnosis utilizing a CNN to identify stenosis (31). In the present study, we aimed to analyze stenosis in MRA using AI. However, AI was considered unsuitable for diagnosis due to instances where the stenotic lesion occluded during the stroke onset; therefore, we employed manual entry. Although this method may miss some cases of stenosis, it was possible to detect stenoses in the ipsilateral carotid artery, internal carotid artery, middle cerebral artery, and basilar artery, and was useful in making a diagnosis. In CE, AI has been developed focusing on the detection of cardiac embolic sources (29). In this study, NT-proBNP was initially considered a potential marker for detecting heart diseases; however, it was deemed unsuitable due to its elevation in patients with chronic kidney disease. Nevertheless, a previous study reported a potential association between NT-proBNP and arrhythmias (32). Further studies with more

TABLE 2 Result for stroke classification AI. Matching table for each stroke type: CE, LAS, SVO/BAD, and ESUS+D in Cohort 1.

		Precision subtype by AI				
		CE	LAS	SVO/BAD	Tro	ESUS+D
Manually assigned stroke subtype	CE	**49**	0	1	1	0
	LAS	0	**29**	0	0	0
	SVO/BAD	0	0	**52**	0	1
	Tro	0	1	1	**4**	3
	ESUS+D	2	5	9	2	**71**

Bold numbers indicate the number of each case correctly answered by the AI. Tro, Trousseau syndrome; ESUS + D, embolic stroke of undetermined sources + cerebral artery dissection; TP, true positives; FN, false negatives; FP, false positives.

TABLE 3 Result for stroke classification AI. Precision rate for each stroke classification.

Stroke classification	Patient number TP + FN	Predicted number of cases TP + FP	Number of correct cases TP	Sensitivity TP/(TP + FN)	Precision TP/(TP + FP)	*F* value
CE	51	51	49	96%	96%	0.96
LAS	29	35	29	100%	83%	0.91
SVO/BAD	53	63	52	98%	83%	0.90
Tro	9	7	4	44%	57%	0.50
ESUS+D	89	75	71	80%	95%	0.87

Tro, Trousseau syndrome; ESUS + D, embolic stroke of undetermined sources + cerebral artery dissection; TP, true positives; FN, false negatives; FP, false positives.

TABLE 4 Result for stroke classification AI. Matching table for each stroke type.

		Precision subtype by AI								
		CE	CE-PFO	LAS	SVO	BAD	Aorta	Tro	Dissec	other
Classification from physician	CE	**49**	0	0	0	1	0	1	0	0
	CE-PFO	0	**0**	0	0	1	1	0	0	5
	LAS	0	0	**29**	0	0	0	0	0	0
	SVO	0	0	0	**25**	0	0	0	0	0
	BAD	0	0	0	0	**27**	1	0	0	0
	Aorto	0	0	2	1	0	**30**	2	0	5
	Tro	0	0	1	0	1	3	**4**	0	0
	Dissec	2	0	2	4	1	4	0	**1**	1
	Other	0	0	1	2	0	10	0	2	**12**

Bolded numbers indicate the number of each case correctly answered by the AI. CE, cardiogenic embolism; CE-PFO, cerebral embolization through a patent foramen ovale (paradoxical embolism); LAS, large artery atherosclerosis; SVO, small vessel occlusion; BAD, branch atheromatous disease; Aorta, aortogenic embolism; Tro, Trousseau; Dissec, cerebral artery dissection; Other, undetermined etiology; TP, true positives; FN, false negatives; FP, false positives.

TABLE 5 Result for stroke classification AI. Precision rate for each stroke classification.

Stroke classification	Patient number TP + FN	Predicted number of cases TP + FP	Number of correct cases TP	Sensitivity TP/(TP + FN)	Precision TP/(TP + FP)	*F* value
CE	51	51	49	96%	96%	0.96
CE-PFO	7	0	0	0%	–	–
LAS	29	35	29	100%	83%	0.91
SVO	25	32	25	100%	78%	0.88
BAD	28	31	27	96%	87%	0.92
Aorto	40	49	30	75%	61%	0.67
Tro	9	7	4	44%	57%	0.50
Dissec	15	3	1	7%	33%	0.11
Other	27	23	12	44%	52%	0.48

CE, cardiogenic embolism; CE-PFO, cerebral embolization through a patent foramen ovale (paradoxical embolism); LAS, large artery atherosclerosis; SVO, small vessel occlusion; BAD, branch atheromatous disease; Aorta, aortogenic embolism; Tro, Trousseau; Dissec, cerebral artery dissection; Other, undetermined etiology; TP, true positives; FN, false negatives; FP, false positives.

cases are needed to establish an association curve between eGFR and NT-proBNP, which would make it a suitable candidate for the development of a stroke diagnosis AI. Furthermore, an algorithm attributed to ESUS has been created (33), which may become a candidate in the future. Initially, D-dimer alone was utilized in this study for Trousseau syndrome, but evaluating it was challenging due to the inclusion of many CE cases within the cutoff line. Incorporating data on active malignancy allowed us to achieve a more defined evaluation. Nonetheless, in real clinical scenarios, there are instances where malignancy is unknown at the time of admission, which may impact the accuracy of the response rate.

Although there are limited reports on TOAST classification of ischemic stroke using AI. Primarily, extraction is done from the electronic health records (EHRs), and Garg et al. (34) extracted

information from 1,091 cases of EHR data and compared it with actual diagnoses using machine learning techniques. The corresponding precision rates obtained were 70.3% for cardioembolic stroke, 65.3% for large artery atherosclerosis (LAA), 62.3% for small vessel occlusion (SVO), and 73.7% for cryptogenic stroke. Additionally, Zhang et al. (35) performed similar analyzes, resulting in precision rates of 53.3% for cardioembolic stroke, 74.5% for LAA, 54.7% for SVO, and 20.0% for cryptogenic stroke. Moreover, Wang et al. (36) conducted an analysis excluding cryptogenic stroke, achieving precision rates of 94.07% for cardioembolic stroke, 76.73% for LAA, and 72.13% for SVO. However, these reports indicate that our developed system exhibits higher diagnostic accuracy. Furthermore, our system has the advantage of performing analysis inclusive of cryptogenic stroke (also known as ESUS) and incorporating image-based analysis. It is believed that the diagnostic accuracy has improved by creating a two-stage system with ischemia detection AI and subtype classification AI. Furthermore, the comparison registry data, although from a single center, is based on a registry that evaluates cryptogenic stroke with transesophageal echocardiography for aortogenic embolism and stroke associated with patent foramen ovale (PFO). It is considered a strength of the AI developed in this study that it can diagnose detailed subtypes of strokes. SCTSS-AI can change the functional prognosis of stroke patients and lead to the equalization of stroke treatment.

In the future, the use of inflammatory markers and cytokines as predictive elements for AI development should be considered. For example, the ligand for CD40 and expression of MCP1 are upregulated in the acute phase of atherothrombotic stroke, which is also associated with vascular events with diabetes (37). Moreover, the association between white blood cell count and blood glucose at onset and mortality during hospitalization, as well as inflammatory markers are potential factors for stroke diagnostic AI (38). Our findings also suggest that hyperglycemia caused by stroke stress was associated with in-hospital mortality, and there may be a relationship with NO activity. Furthermore, an association between peripheral vasoreactivity index and endothelial function has been reported in the LAS (39). Arterial stiffness indices, such as augmentation index and pulse wave velocity, have been shown to be higher in LAS patients, and arterial stiffness indices at onset may also be useful for the establishment of AI diagnosis in the future.

Our study has several limitations. First, it was a single-center, medical record-based, retrospective study. Second, unlike other AI studies, only a few hundred cases were needed for constructing the AI. Although this was a strength of the study, the small number of cases may also be a limitation. Further studies with more cases may be able to distinguish CE-PFO features and SVO from BAD with only an initial MRI. Third, the features required to distinguish ESUS+D were limited; therefore, the number of cases and characteristics of ESUS+D-determined etiology cases require clarification. Moreover, the diagnosis of cerebral artery dissection requires the collection of vertebral/basilar artery MRA findings. In the present study, the number of cases used for training the AI was insufficient for learning to predict the vascular morphology in areas where no vessels were captured on MRA. Further case collection and prospective randomized studies are needed to address these uncertainties. Fourth, this study had technical limitations. Initially, our proposed method constructed the AI model using 231 cases, which was insufficient to cover all stroke variations. While increasing case numbers may improve our model's performance, it also introduces exceptions that our model may not classify accurately. This inherent challenge in AI construction is unavoidable. Continuously expanding the dataset and re-building the AI is necessary to yield benefits for patients. Further studies using novel AI models with higher speed are needed to identify exceptions that our model is unable to classify to optimize model structure and hyperparameters.

In conclusion, AI in stroke imaging has the potential to revolutionize stroke diagnosis and patient management. Diagnosis of stroke using machine learning methods could be especially useful for health care providers who are not familiar with stroke imaging, such as general practitioners and paramedics, and to speed up treatment decisions. This study, which achieved high accuracy in detecting strokes and classifying their vascular regions, may contribute to the automatic detection of strokes, enabling physicians to make quick and

TABLE 6 Result for stroke classification AI. Matching table for each stroke type; CE, LAS, SVO/BAD, and ESUS+D in Cohort 2.

		Precision subtype by AI				
		CE	LAS	SVO/BAD	Tro	ESUS+D
Manually assigned stroke subtype	CE	**44**	2	4	0	2
	LAS	0	**22**	2	0	1
	SVO/BAD	0	0	**30**	0	2
	Tro	0	0	0	**6**	3
	ESUS+D	2	0	2	0	**28**

Bold numbers indicate the number of each case correctly answered by the AI. Tro, Trousseau syndrome; ESUS + D, embolic stroke of undetermined sources + cerebral artery dissection; TP, true positives; FN, false negatives; FP, false positives.

TABLE 7 Result for stroke classification AI. Precision rate for each stroke classification.

Stroke classification	Patient number TP + FN	Predicted number of cases TP + FP	Number of correct cases TP	Sensitivity TP/(TP + FN)	Precision TP/(TP + FP)	*F* value
CE	52	46	44	85%	96%	0.90
LAS	25	24	22	88%	92%	0.90
SVO/BAD	33	38	30	91%	79%	0.85
Tro	9	7	6	67%	86%	0.75
ESUS+D	32	36	28	88%	78%	0.82

Tro, Trousseau syndrome; ESUS + D, embolic stroke of undetermined sources + cerebral artery dissection; TP, true positives; FN, false negatives; FP, false positives.

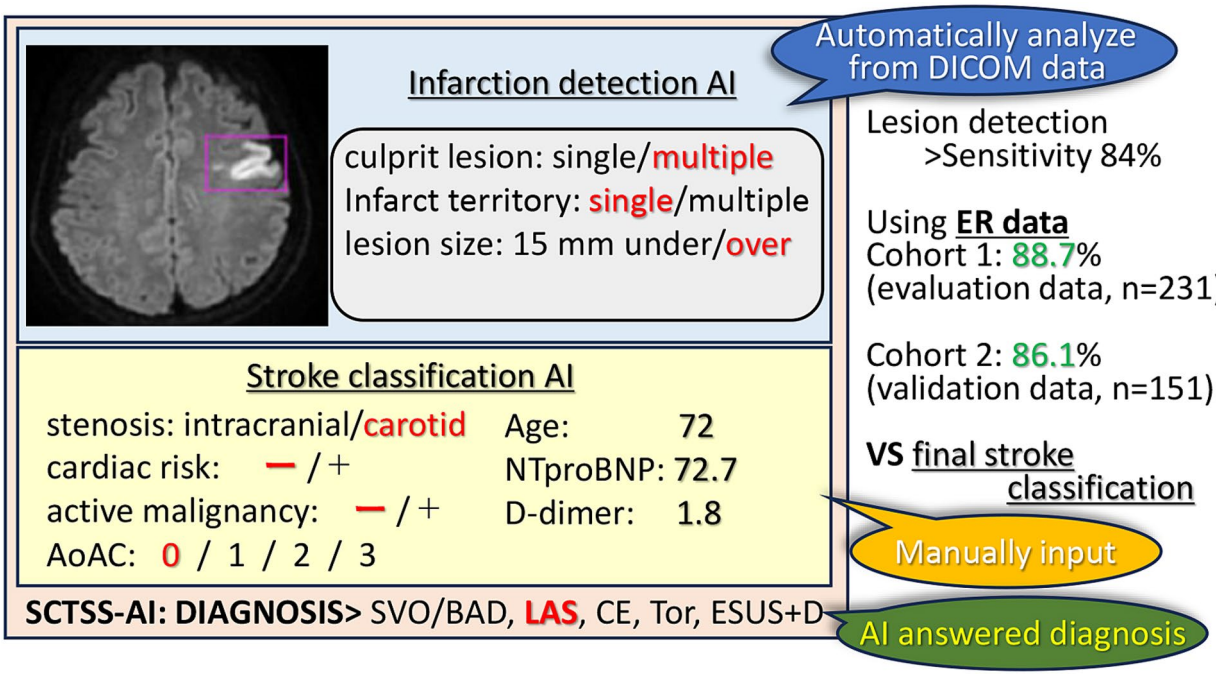


FIGURE 4
SCTSS-AI concept diagram. The SCTSS-AI identifies and classifies the infarct features (location, number, size) from the DWI on the DICOM data. Although manual input is required, SCTSS-AI automatically determines the type of disease by entering various parameters. Compared to the final stroke classification, the SCTSS-AI provides a definitive diagnosis with an accuracy of more than 86% using data obtained only from the emergency department.

appropriate treatment decisions (Figure 4). Since the AI was created using only factors that are known in the emergency department, we were able to establish an AI that is directly related to clinical practice. Furthermore, our findings suggest that additional tests, such as transesophageal echocardiography and Holter EEG analysis, should be performed if the patient is classified as ESUS+D.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Human Ethics Review Committee of Juntendo University School of Medicine (E22-0028). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. YU: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Validation, Writing – review & editing. KY: Methodology, Writing – review & editing. KH: Conceptualization, Investigation, Methodology, Writing – review & editing. CK: Investigation, Resources, Writing – review & editing. NK: Investigation, Writing – review & editing. YI: Investigation, Writing – review & editing. KOk: Investigation, Writing – review & editing. SM: Investigation, Writing – review & editing. DT: Writing – review & editing. KOn: Writing – review & editing. MA: Writing – original draft, Writing – review & editing. KM: Writing – review & editing. NH: Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported partly by Grants-in-Aid from the Foundation of Strategic Research Projects in Private Universities from the Ministry of Education, Culture, Sports, Science, and Technology, and Ohara Pharmaceutical Co., Ltd.

Conflict of interest

NK, YI, and SM were employed by HACARUS INC. KOk and DT are employed by HACARUS INC. KOn and MA are employed by

Ohara Pharmaceutical Co., Ltd. KM is a board member of Ohara Pharmaceutical Co., Ltd. MA is employed by PARKINSON Laboratories Co., Ltd. KM is a board member of PARKINSON Laboratories Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Bonita R. Epidemiology of stroke. *Lancet*. (1992) 339:342–4. doi: 10.1016/0140-6736(92)91658-U

2. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet*. (2006) 367:1747–57. doi: 10.1016/S0140-6736(06)68770-9

3. Smith SC Jr. Reducing the global burden of ischemic heart disease and stroke: a challenge for the cardiovascular community and the United Nations. *Circulation*. (2011) 124:278–9. doi: 10.1161/CIRCULATIONAHA.111.040170

4. Feigin VL, Lawes CM, Bennett DA, Anderson CS. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol*. (2003) 2:43–53. doi: 10.1016/S1474-4422(03)00266-7

5. Dirnagl U, Iadecola C, Moskowitz MA. Pathobiology of ischaemic stroke: an integrated view. *Trends Neurosci*. (1999) 22:391–7. doi: 10.1016/S0166-2236(99)01401-0

6. Latchaw RE, Alberts MJ, Lev MH, Connors JJ, Harbaugh RE, Higashida RT, et al. Recommendations for imaging of acute ischemic stroke: a scientific statement from the American Heart Association. *Stroke*. (2009) 40:3646–78. doi: 10.1161/STROKEAHA.108.192616

7. Meyers PM, Schumacher HC, Connolly ES Jr, Heyer EJ, Gray WA, Higashida RT. Current status of endovascular stroke treatment. *Circulation*. (2011) 123:2591–601. doi: 10.1161/CIRCULATIONAHA.110.971564

8. Fiez JA, Damasio H, Grabowski TJ. Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. *Hum Brain Mapp*. (2000) 9:192–211. doi: 10.1002/(SICI)1097-0193(200004)9:4<192::AID-HBM2>3.0.CO;2-Y

9. Ashton EA, Takahashi C, Berg MJ, Goodman A, Totterman S, Ekholm S. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. *J Magn Reson Imaging*. (2003) 17:300–8. doi: 10.1002/jmri.10258

10. Filippi M, Horsfield MA, Bressi S, Martinelli V, Baratti C, Reganati P, et al. Intra- and inter-observer agreement of brain MRI lesion volume measurements in multiple sclerosis. A comparison of techniques. *Brain*. (1995) 118:1593–600. doi: 10.1093/brain/118.6.1593

11. Wilke M, de Haan B, Juenger H, Karnath HO. Manual, semi-automated, and automated delineation of chronic brain lesions: a comparison of methods. *NeuroImage*. (2011) 56:2038–46. doi: 10.1016/j.neuroimage.2011.04.014

12. Lansberg MG, Albers GW, Beaulieu C, Marks MP. Comparison of diffusion-weighted MRI and CT in acute stroke. *Neurology*. (2000) 54:1557–61. doi: 10.1212/WNL.54.8.1557

13. Ikki Y, Yamada M, Sekine M. Regional disparity of certified teaching hospitals on physicians' workload and wages, and popularity among medical students in Japan. *Environ Health Prev Med*. (2021) 26:75. doi: 10.1186/s12199-021-00997-3

14. Miyamoto S, Ogasawara K, Kuroda S, Itabashi R, Toyoda K, Itoh Y, et al. Japan stroke society guideline 2021 for the treatment of stroke. *Int J Stroke*. (2022) 17:1039–49. doi: 10.1177/17474930221090347

15. Kato D, Ryu H, Matsumoto T, Abe K, Kaneko M, Ko M, et al. Building primary care in Japan: literature review. *J Gen Fam Med*. (2019) 20:170–9. doi: 10.1002/jgf2.252

16. Parr E, Ferdinand P, Roffe C. Management of Acute Stroke in the older person. *Geriatrics (Basel)*. (2017) 2:27. doi: 10.3390/geriatrics2030027

17. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of org 10172 in acute stroke treatment. *Stroke*. (1993) 24:35–41. doi: 10.1161/01.STR.24.1.35

18. Kasner SE. Clinical interpretation and use of stroke scales. *Lancet Neurol*. (2006) 5:603–12. doi: 10.1016/S1474-4422(06)70495-1

19. Senin P. *Dynamic time warping algorithm review*. Manoa Honolulu: Information and Computer Science Department University of Hawaii (2008). 40 p.

20. Shimada Y, Ueno Y, Tanaka Y, Okuzumi A, Miyamoto N, Yamashiro K, et al. Aging, aortic arch calcification, and multiple brain infarcts are associated with aortogenic brain embolism. *Cerebrovasc Dis*. (2013) 35:282–90. doi: 10.1159/000347073

21. Lee LJ, Kidwell CS, Alger J, Starkman S, Saver JL. Impact on stroke subtype diagnosis of early diffusion-weighted magnetic resonance imaging and magnetic resonance angiography. *Stroke*. (2000) 31:1081–9. doi: 10.1161/01.STR.31.5.1081

22. Sammut C, Webb GI. Leave-One-Out Cross-Validation In: C Sammut and GI Webb, editors. *Encyclopedia of machine learning*. New York City: Springer Publishing (2010). 600–1.

23. Yedavalli VS, Tong E, Martin D, Yeom KW, Forkert ND. Artificial intelligence in stroke imaging: current and future perspectives. *Clin Imaging*. (2021) 69:246–54. doi: 10.1016/j.clinimag.2020.09.005

24. Subudhi A, Sahoo S, Biswal P, Sabut S. Segmentation and classification of ischemic stroke using optimized features in brain MRI. *Biomed Eng - Appl Basis Commun*. (2018) 30:1850011. doi: 10.4015/S1016237218500114

25. Cetinoglu YK, Koska IO, Uluc ME, Gelal MF. Detection and vascular territorial classification of stroke on diffusion-weighted MRI by deep learning. *Eur J Radiol*. (2021) 145:110050. doi: 10.1016/j.ejrad.2021.110050

26. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Commun ACM*. (2017) 60:84–90. doi: 10.1145/3065386

27. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE conference on computer vision and pattern recognition.

28. Tan M, Le Q (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. International conference on machine learning.

29. Miceli G, Basso MG, Rizzo G, Pintus C, Cocciola E, Pennacchio AR, et al. Artificial intelligence in acute ischemic stroke subtypes according to Toast classification: a comprehensive narrative review. *Biomedicine*. (2023) 11:1138. doi: 10.3390/biomedicines11041138

30. Macharzina RR, Kocher S, Messe SR, Rutkowski T, Hoffmann F, Vogt M, et al. 4-dimensionally guided 3-dimensional color-Doppler ultrasonography quantifies carotid artery stenosis with high reproducibility and accuracy. *JACC Cardiovasc Imaging*. (2018) 11:386–96. doi: 10.1016/j.jcmg.2017.02.018

31. Bonanno L, Marino S, Bramanti P, Sottile F. Validation of a computer-aided diagnosis system for the automatic identification of carotid atherosclerosis. *Ultrasound Med Biol*. (2015) 41:509–16. doi: 10.1016/j.ultrasmedbio.2014.09.004

32. Ueno Y, Miyamoto N, Hira K, Doijiri R, Yamazaki H, Sonoda K, et al. Left atrial appendage flow velocity predicts occult atrial fibrillation in cryptogenic stroke: a CRYPTON-ICM registry. *J Neurol*. (2023) 270:5878–88. doi: 10.1007/s00415-023-11942-5

33. Ntaios G, Weng SF, Perlepe K, Akyea R, Condon L, Lambrou D, et al. Data-driven machine-learning analysis of potential embolic sources in embolic stroke of undetermined source. *Eur J Neurol*. (2021) 28:192–201. doi: 10.1111/ene.14524

34. Garg R, Oh E, Naidech A, Kording K, Prabhakaran S. Automating ischemic stroke subtype classification using machine learning and natural language processing. *J Stroke Cerebrovasc Dis*. (2019) 28:2045–51. doi: 10.1016/j.jstrokecerebrovasdis.2019.02.004

35. Zhang S, Wang J, Pei L, Liu K, Gao Y, Fang H, et al. Interpretable CNN for ischemic stroke subtype classification with active model adaptation. *BMC Med Inform Decis Mak*. (2022) 22:3. doi: 10.1186/s12911-021-01721-5

36. Wang J, Gong X, Chen H, Zhong W, Chen Y, Zhou Y, et al. Causative classification of ischemic stroke by the machine learning algorithm random forests. *Front Aging Neurosci*. (2022) 14:788637. doi: 10.3389/fnagi.2022.788637

37. Davi G, Tuttolomondo A, Santilli F, Basili S, Ferrante E, Di Raimondo D, et al. CD40 ligand and MCP-1 as predictors of cardiovascular events in diabetic patients with stroke. *J Atheroscler Thromb*. (2009) 16:707–13. doi: 10.5551/jat.1537

38. Tuttolomondo A, Pedone C, Pinto A, Di Raimondo D, Fernandez P, Di Sciacca R, et al. Predictors of outcome in acute ischemic cerebrovascular syndromes: the GIFA study. *Int J Cardiol*. (2008) 125:391–6. doi: 10.1016/j.ijcard.2007.03.109

39. Tuttolomondo A, Casuccio A, Della Corte V, Maida C, Pecoraro R, Di Raimondo D, et al. Endothelial function and arterial stiffness indexes in subjects with acute ischemic stroke: relationship with TOAST subtype. *Atherosclerosis*. (2017) 256:94–9. doi: 10.1016/j.atherosclerosis.2016.10.044

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 14 December 2023
DOI 10.3389/fneur.2023.1290117

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Elisa Gouvêa Bogossian,
Université libre de Bruxelles, Belgium
Jaime Daniel Mondragón,
San Diego State University, United States

*CORRESPONDENCE
Jun Lyu
✉ lyujun2020@jnu.edu.cn

†These authors have contributed equally to this work

RECEIVED 11 September 2023
ACCEPTED 30 November 2023
PUBLISHED 14 December 2023

CITATION
Guo J, Cheng H, Wang Z, Qiao M, Li J and Lyu J (2023) Factor analysis based on SHapley Additive exPlanations for sepsis-associated encephalopathy in ICU mortality prediction using XGBoost — a retrospective study based on two large database.
Front. Neurol. 14:1290117.
doi: 10.3389/fneur.2023.1290117

COPYRIGHT
© 2023 Guo, Cheng, Wang, Qiao, Li and Lyu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Factor analysis based on SHapley Additive exPlanations for sepsis-associated encephalopathy in ICU mortality prediction using XGBoost — a retrospective study based on two large database

Jiayu Guo[1,2†], Hongtao Cheng[1,3†], Zicheng Wang[1], Mengmeng Qiao[1,2], Jing Li[1,2] and Jun Lyu[1,4*]

[1]Department of Clinical Research, The First Affiliated Hospital of Jinan University, Guangzhou, China, [2]School of Public Health, Shannxi University of Chinese Medicine, Xianyang, China, [3]School of Nursing, Jinan University, Guangzhou, Guangdong, China, [4]Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization, Guangzhou, Guangdong, China

Objective: Sepsis-associated encephalopathy (SAE) is strongly linked to a high mortality risk, and frequently occurs in conjunction with the acute and late phases of sepsis. The objective of this study was to construct and verify a predictive model for mortality in ICU-dwelling patients with SAE.

Methods: The study selected 7,576 patients with SAE from the MIMIC-IV database according to the inclusion criteria and randomly divided them into training (n = 5,303, 70%) and internal validation (n = 2,273, 30%) sets. According to the same criteria, 1,573 patients from the eICU-CRD database were included as an external test set. Independent risk factors for ICU mortality were identified using Extreme Gradient Boosting (XGBoost) software, and prediction models were constructed and verified using the validation set. The receiver operating characteristic (ROC) and the area under the ROC curve (AUC) were used to evaluate the discrimination ability of the model. The SHapley Additive exPlanations (SHAP) approach was applied to determine the Shapley values for specific patients, account for the effects of factors attributed to the model, and examine how specific traits affect the output of the model.

Results: The survival rate of patients with SAE in the MIMIC-IV database was 88.6% and that of 1,573 patients in the eICU-CRD database was 89.1%. The ROC of the XGBoost model indicated good discrimination. The AUCs for the training, test, and validation sets were 0.908, 0.898, and 0.778, respectively. The impact of each parameter on the XGBoost model was depicted using a SHAP plot, covering both positive (acute physiology score III, vasopressin, age, red blood cell distribution width, partial thromboplastin time, and norepinephrine) and negative (Glasgow Coma Scale) ones.

Conclusion: A prediction model developed using XGBoost can accurately predict the ICU mortality of patients with SAE. The SHAP approach can enhance the interpretability of the machine-learning model and support clinical decision-making.

KEYWORDS

sepsis-associated encephalopathy (SAE), XGBoost, SHAP (SHapley Additive exPlanations), ICU mortality, eICU-CRD, MIMIC-IV

Introduction

Sepsis, a syndrome caused by dysfunction of organs including the central nervous system (CNS), heart, and lungs (1–3) due to dysregulation of the host response to infection, is the most common cause of death in intensive care unit patients worldwide. One manifestation of sepsis-induced cerebral dysfunction is sepsis-associated encephalopathy (SAE), which is defined as diffuse cerebral dysfunction secondary to organic infection in the absence of an obvious central nervous system infection (4).

The pathophysiology of SAE is intricate, arising from a convergence of inflammatory and non-inflammatory processes impacting various categories of cerebral cells. Significant mechanisms encompass heightened microglial activation, disruption of the blood–brain barrier (BBB), and the perpetuation of an extended inflammatory reaction (5). Upon the initial emergence of sepsis, an inordinate immune-inflammatory response is incited, setting in motion the infiltration of inflammatory mediators into cerebral tissue, thereby activating microglial cells. This activation gives rise to the establishment of a cytotoxic milieu, instigating the release of reactive oxygen species, nitric oxide (6), and glutamate, as a countermeasure against sepsis. Nevertheless, the CNS is notably vulnerable to neurotoxic agents such as free radicals, inflammatory mediators, and intravascular proteins, thus precipitating a malfunction in the BBB (7). The relentless activation of microglia perpetuates a deleterious cycle, culminating in aberrant neuronal performance and cellular demise, thereby exacerbating BBB impairment and the progression of SAE. In addition to this, sepsis damages the hippocampus, cortex, cerebellum and brainstem of the brain. Sepsis-driven brain damage occurs in a diffuse form and is strongly associated with cognitive impairment.

Clinicians must exclude primary CNS disorders, sedation-related cognitive disorders, metabolic encephalopathies, and poisonings before diagnosing SAE on the basis of cognitive and neuropsychiatric deficits, manifestations of delirium, or a Glasgow Coma Scale (GCS) score of less than 15 (8). Globally, up to 50% of intensive care unit (ICU) patients present with SAE during sepsis (4, 9), which tends to increase the length of stay and mortality of septic patients in the ICU (10). The current lack of specific treatment options and insufficient understanding of the underlying mechanisms of SAE are the most common causes of poor prognosis in sepsis. Therefore, the aim of this study was to investigate the independent risk factors for ICU death in patients with SAE and to develop a predictive model to quantify the likelihood of ICU death in patients with SAE.

Materials and methods

Data source

Data for this study were obtained from the MIMIC-IV and eICU-CRD databases, with the former being a multiparametric, structured single-center critical-care database published in 2003 that includes clinically available data on more than 380,000 patients during 2008–2019. There was no requirement to obtain permission from individual patients or ethical approval statements because the initiative had no impact on clinical care and none of the patients in the database could be identified (11). Our study also followed the guidelines of the Declaration of Helsinki and Transparent Reporting of a Multivariate Prediction Model for Individual Prognosis or Diagnosis (12).

The eICU-CRD database contains data from the ICU wards of numerous hospitals in the US. It contains routine data on 200,859 patients obtained from more than 300 hospitals in the US during 2014 and 2015 (13). No specific patient permission was needed because both databases use anonymous health data.

Patient population

Presently, there exists a deficiency of precise diagnostic modalities for SAE. Clinical diagnosis relies on exclusion and necessitates discrimination from central nervous system infections, metabolic encephalopathy (a widespread yet potentially reversible cerebral dysfunction arising from metabolic or toxic origins), excessive sedative ingestion, and withdrawal manifestations with the potential to impact sensory faculties.

Patients with a Sequential Organ Failure Assessment (SOFA) score ≥ 2 based on the Sepsis-3 classification and a GCS score < 15 or delirium on the day before admission to the ICU were considered SAE patients. The exclusion criteria were (1) presence of primary brain injury, (2) psychiatric disorders and neurological diseases, (3) metabolic, hepatic, hypertensive, or toxic encephalopathy, (4) severe electrolyte disturbance or deglycation, (5) patients who were intubated, given analgesics, and sedated at the time of admission, (6) long-term alcohol or drug abuse, or (7) an ICU stay of <24 h. Figure 1 depicts the flow chart for case inclusion.

Observation indicators

This study used Structured Query Language to extract the following basic information of patients from both databases: age, gender, and mean values of vital signs at the time of first ICU admission, including heart rate, respiratory rate, and body temperature. From the time of ICU admission, the first laboratory data included ghrelin, lymphocytes, eosinophils, neutrophils, monocytes, hemoglobin, urea nitrogen, platelets, creatinine, glucose, blood urea nitrogen (BUN), hematocrit (HCT), partial thromboplastin time (PTT), white blood cell count (WBC), normalized ratio (INR), anion gap (AG), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), Mean Corpuscular Volume (MCV) and red blood cell distribution width (RDW), Severity

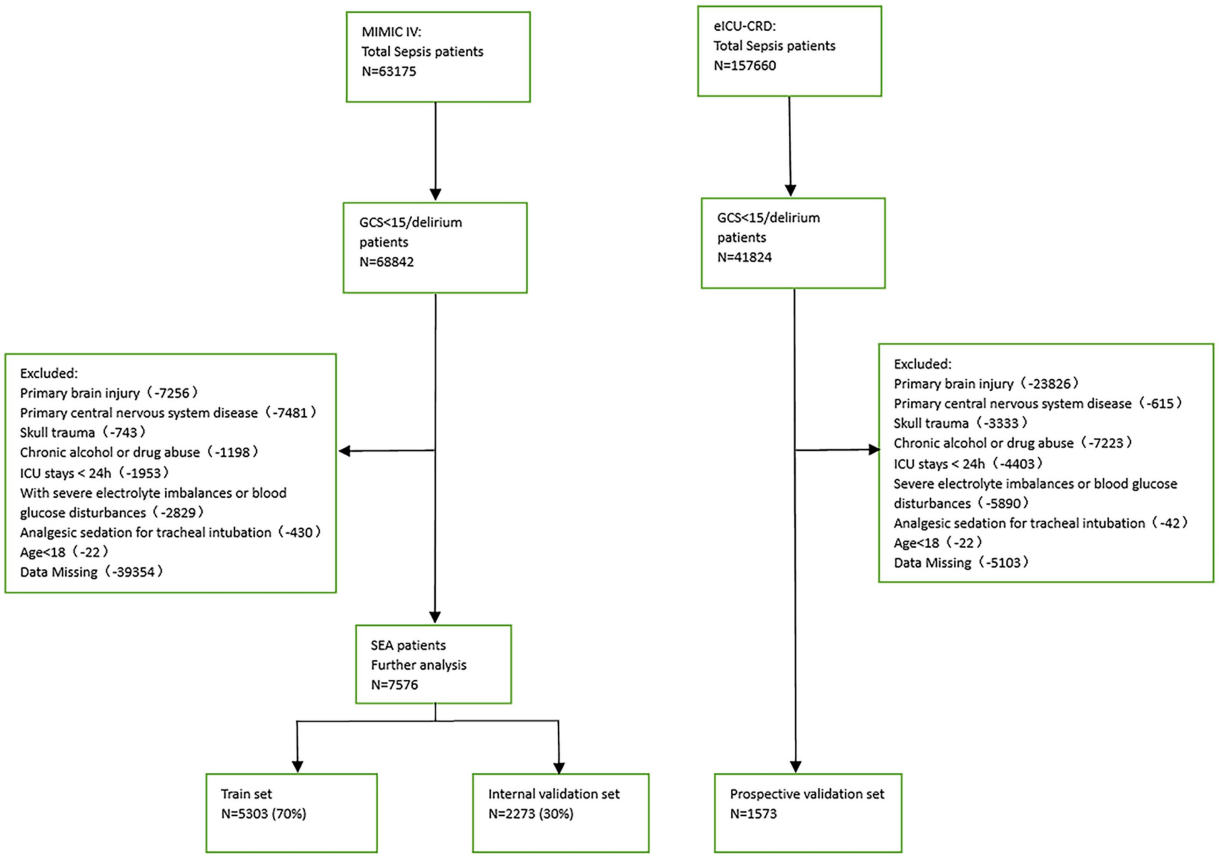


FIGURE 1
Flowchart of patient cohorts.

was determined using the acute physiology score III (APSIII), SOFA score, GCS score, and comorbidity occurrence.

Statistical analysis

Statistical processing was performed using R software (version 4.2.2). For continuous variables, values are expressed as standard deviation or median of interquartile range (IQR); for categorical variables, values are expressed as total (%). Comparisons of continuous variables were made using the *t* test or Wilcoxon rank sum test, and comparisons of proportions were made using the χ 2 test or Fisher exact test. After the variables were identified by Extreme Gradient Boosting (XGBoost), we used these included clinical and laboratory variables to construct a prediction model for in-ICU mortality in SAE patients based on the XGBoost algorithm. XGBoost is an improved algorithm based on gradient boosting decision trees that efficiently constructs boosted trees and runs them in parallel (14). The core of the algorithm is to optimize the value of the objective function (15). In model development and comparison, we use a 5-fold cross-validation approach, which provides a more stable and reliable way to measure the performance of the model.

The prediction model was trained and internally validated using training and test sets randomized at a ratio of 7:3. The performance of the prediction model was externally validated using the identical data of patients with SAE from the eICU-CRD database. The predictive values of different models were analyzed using the receiver operating characteristic (ROC) and area under the ROC curve (AUC), with the latter allowing quantitative differentiation of column line graphs. In the XGBoost analysis, qualitative data were converted to numerical data, and "yes" and "no" were converted to "1" and "0," respectively.

An aesthetically pleasing additional interpretation method, the SHapley Additive exPlanations (SHAP), was used in XGBoost to increase the readability of the model. SHAP is a technique used to explain the output of any machine-learning model (16). A SHAP summary plot was used to present the effect of the characteristics attributed to the model. Colors in the scatter plot intuitively represent the correlation between the characteristic value and the anticipated probability. The importance of specific features and their impacts on the output of the model were examined using the SHAP dependence plot. A SHAP force plot was used to illustrate how important characteristics affect use of the final model across all patients.

Results

Baseline patient characteristics

This study used 1,573 patients from the eICU-CRD database as the external test set, and 7,576 patients from the MIMIC-IV database were randomly split into a training ($n = 5,303$, 70%) and an internal test ($n = 2,273$, 30%) set.

Among the 7,576 patients within the MIMIC-IV database afflicted by SAE, 6709 (88.6%) experienced survival, while 867 (11.4%) succumbed prior to ICU discharge. In the case of the 1,573 patients drawn from the eICU-CRD database, 1402 (89.1%) survived, and 171 (10.9%) met their demise. Across the training, internal test, and external test sets, the majority of patients were aged over 65 years (with mean ages of 69, 70, and 66 years, respectively) and were predominantly male (comprising 59.2, 59.6, and 57.8% of the respective cohorts). The average length of ICU stay was 3.38 days for the training set, 3.29 days for the internal test set, and 3.99 days for the external test set. Additionally, many of these patients presented with comorbidities, including congestive heart failure (35.9, 36.6, and 7.3%), COPD (28.1, 25.9, and 1.1%), diabetes (9.4, 9.2, and 0.8%), renal failure (25.8, 25.6, and 9.4%), or liver disease (11.5, 13.0, and 1.4%).

The demographic profiles and fundamental patient information pertaining to the training and test sets are delineated in Tables 1, 2. In Table 3, we present the foundational attributes of the study cohort sourced from the MIMIC-IV database, stratified by distinct outcomes. Notably, the average age of patients afflicted by SAE was notably higher in the deceased group in comparison to the survivor group. Furthermore, the incidence rates of myocardial infarction, peripheral vascular disease, dementia, diabetes, sclerosis, and liver disease exhibited variations between these two cohorts. Conversely, no statistically significant disparities were observed between the two groups concerning myocardial infarction, peripheral vascular disease, dementia, diabetes, sodium levels, and MCH. Turning our attention to the baseline attributes of the subjects derived from the eICU-CRD database for different outcomes, these details are summarized in Table 4. Significant differences among groups were discerned in variables such as sedative usage, analgesic administration, vasopressin and norepinephrine dosages, GCS score, SOFA score, lactate levels, creatinine values, bicarbonate levels, BUN, PTT, INR, AG, MCHC, RDW, respiratory rate, and body temperature.

Feature selection

The XGBoost algorithm identified APSIII, vasopressin, GCS score, PTT, norepinephrine, age, RDW, and length of ICU stay as independent predictors of SAE. Figure 2A presents the importance of each factor influencing SAE. APSIII had the highest score, indicating that determining severity in patients was the most relevant and important factor. Smaller APSIII values indicate a lower output from the model. The GCS score had the smallest effect on the model. Figure 2B presents the SHAP summary plot, which reflects the influence of each factor using the SHAP value in XGBoost and whether they had a positive or negative effect The SHAP plot illustrates the influence of each parameter on the XGBoost model, including the positive (APSIII, vasopressin, age, RDW, PTT, and norepinephrine) and negative (GCS) effects.

Each of the eight factors is represented by a SHAP dependence plot in Figure 3, which illustrates how different characteristics influenced the XGBoost model results. Positive SHAP values for specific factors represent an elevated mortality risk. We found that mortality was correlated with higher APSIII, age, RDW, and PTT, and a lower GCS score. Both longer and shorter ICU stays were associated with lower survival rates. Patients who receive vasopressin and norepinephrine may experience higher mortality rates.

When GCS scores were low, the SHAP interaction values of APSIII with the GCS score decreased as APSIII increased (Figure 4A). The interaction effect of APSIII with norepinephrine and vasopressin (Figures 4B,C) did not seem to be affected by differences in norepinephrine or vasopressin use. Samples with the highest SHAP values for death in the ICU were often accompanied by vasopressin use and a shorter ICU stay. When the GCS score was higher, the value of the interaction between time and GCS score decreased as the length of ICU stay increased (Figure 4D). The interaction effect of time in ICU with norepinephrine (Figure 4E) did not appear to be affected by the use of norepinephrine. The value of the interaction between time and vasopressin use decreased as the length of ICU stay increased (Figure 4F). When the GCS score was high, the interaction value between age and GCS score decreased as age increased (Figure 4G), and the interaction value between age and GCS score decreased to a negative value at 73 years old. When norepinephrine was used, the value of the interaction between age and norepinephrine increased with age (Figure 4H) and became positive at 73 years old. The SHAP interaction values for age with vasopressin use also increased with age (Figure 4I).

The ultimate output was obtained as the sum of the attributions from each predictor, as seen in the SHAP force plot (Figure 5), which displays these SHAP values stacked for each observation.

Discrimination ability

The ROC was used to evaluate the discrimination ability of the model. The XGBoost model test, internal validation, and external validation sets had AUC values of 0.908, 0.898, and 0.778, respectively (Figure 6).

Discussion

SAE represents a multifaceted encephalopathy, signifying a widespread cerebral impairment stemming from sepsis. It manifests with manifestations such as delirium, coma, cognitive deficits encompassing the loss of learning and memory, and the occurrence of seizures. The pathophysiology of SAE remains partially elucidated. While advances in sepsis research and treatment have lately yielded enhanced prognostic outcomes, the mortality rate of SAE remains disheartening. The identification of risk factors is imperative in grasping the prognosis of SAE.

An XGBoost model was constructed and validated in this study to predict ICU mortality in patients with SAE. The importance analysis of the factors in the XGBoost model suggested that APSIII, vasopressin, age, length of ICU stay, RDW, norepinephrine, PTT, and GCS score are strong predictors of SAE. In this study, the XGBoost prediction model was obtained from the AUC results and had good predictive power. SHAP also offered credible visual interpretation of the predictions, encompassing both positive and negative impacts. In this study we not only calculated values of general parameters for predicting the probability of death in the ICU, but also presented a visual

TABLE 1 Research subject base information form (internal validation).

	Total	Training cohorts	Internal validation cohorts	*p* value
N	7,576	5,303	2,273	
Death = No/Yes (%)	6709/867 (88.6/11.4)	4690/613 (88.4/11.6)	2019/254 (88.8/11.2)	0.658
ICU stay time, days (median [IQR])	3.33 [2.00, 6.29]	3.38 [2.00, 6.33]	3.29 [1.96, 6.13]	0.363
Age (median [IQR])	70.00 [58.00, 79.00]	69.00 [58.00, 79.00]	70.00 [59.00, 79.00]	0.706
Gender = Male/Female (%)	4497/3079 (59.4/40.6)	3142/2161 (59.2/40.8)	1355/918 (59.6/40.4)	0.787
Medical treatments, n (%)				
Sedatives = No/Yes	1292/6284 (17.1/82.9)	900/4403 (17.0/83.0)	392/1881 (17.2/82.8)	0.797
Analgesic = No/Yes	729/6847 (9.6/90.4)	492/4811 (9.3/90.7)	237/2036 (10.4/89.6)	0.131
Antibiotic = No/Yes	988/6588 (13.0/87.0)	689/4614 (13.0/87.0)	299/1974 (13.2/86.8)	0.877
Vasopressin = No/Yes	6642/934 (87.7/12.3)	4634/669 (87.4/12.6)	2008/265 (88.3/11.7)	0.262
Comorbidity, n (%)				
Myocardial infarct = No/Yes	5997/1579 (79.2/20.8)	4193/1110 (79.1/20.9)	1804/469 (79.4/20.6)	0.793
Congestive heart failure = No/Yes	4837/2739 (63.8/36.2)	3397/1906 (64.1/35.9)	1440/833 (63.4/36.6)	0.576
Peripheral vascular disease = No/Yes	6397/1179 (84.4/15.6)	4468/835 (84.3/15.7)	1929/344 (84.9/15.1)	0.523
Dementia = No/Yes	7353/223 (97.1/2.9)	5151/152 (97.1/2.9)	2202/71 (96.9/3.1)	0.594
COPD = No/Yes	5499/2077 (72.6/27.4)	3814/1489 (71.9/28.1)	1685/588 (74.1/25.9)	0.051
Liver disease = No/Yes	6671/905 (88.1/11.9)	4694/609 (88.5/11.5)	1977/296 (87.0/13.0)	0.064
Diabetes = No/Yes	6871/705 (90.7/9.3)	4807/496 (90.6/9.4)	2064/209 (90.8/9.2)	0.862
Renal disease = No/Yes	5627/1949 (74.3/25.7)	3935/1368 (74.2/25.8)	1692/581 (74.4/25.6)	0.852
Severe score, median (IQR)				
SOFA	7.00 [5.00, 10.00]	7.00 [5.00, 9.00]	7.00 [5.00, 10.00]	0.12
GCS	13.00 [8.00, 14.00]	13.00 [8.00, 14.00]	13.00 [8.00, 14.00]	0.054
APSIII	54.00 [39.00, 75.00]	54.00 [39.00, 75.00]	54.00 [39.00, 77.00]	0.298
Laboratory tests, median (IQR)				
Lactate (mmol/L)	1.80 [1.30, 2.60]	1.80 [1.30, 2.60]	1.80 [1.30, 2.55]	0.4
Glucose (mg/dl)	128.50 [113.50, 144.78]	128.25 [113.00, 145.07]	129.10 [114.94, 144.25]	0.303
Creatinine (mg/dl)	1.05 [0.75, 1.65]	1.05 [0.75, 1.65]	1.05 [0.75, 1.65]	0.854
BUN (K/uL)	21.00 [14.50, 35.50]	21.00 [14.50, 35.00]	21.50 [14.50, 37.00]	0.064
Platelets (K/ul)	178.50 [129.50, 248.00]	180.00 [130.00, 249.00]	176.50 [127.50, 246.00]	0.128
Potassium (K/ul)	4.25 [3.90, 4.65]	4.25 [3.90, 4.65]	4.25 [3.95, 4.65]	0.045
Sodium (K/ul)	138.50 [136.00, 141.00]	138.50 [136.00, 141.00]	138.50 [136.00, 141.00]	0.539
Bicarbonate (meq/L)	23.00 [20.50, 25.50]	23.00 [20.50, 25.50]	23.00 [20.50, 25.00]	0.767
Calcium (mg/dl)	8.20 [7.80, 8.70]	8.20 [7.80, 8.70]	8.20 [7.80, 8.65]	0.93
Chloride (mmol/L)	105.00 [101.00, 108.00]	105.00 [101.00, 108.00]	105.00 [101.00, 108.00]	0.416
HCT (%)	10.10 [8.90, 11.55]	10.10 [8.85, 11.50]	10.15 [8.95, 11.65]	0.143
PTT (s)	32.55 [28.05, 41.66]	32.60 [28.10, 41.90]	32.50 [28.00, 40.90]	0.329
WBC (K/uL)	12.10 [8.85, 16.05]	12.10 [8.85, 16.10]	12.00 [8.80, 15.90]	0.529
INR	1.30 [1.15, 1.55]	1.30 [1.15, 1.55]	1.30 [1.15, 1.55]	0.938
AG	14.00 [12.00, 16.50]	14.00 [12.00, 16.50]	13.50 [11.50, 16.50]	0.654
MCH (pg)	30.00 [28.60, 31.30]	30.00 [28.60, 31.30]	30.10 [28.60, 31.30]	0.783
MCHC (g/dL)	32.80 [31.60, 33.90]	32.80 [31.60, 33.90]	32.80 [31.70, 33.80]	0.899
MCV (fL)	91.00 [87.00, 95.00]	91.00 [87.00, 95.00]	91.00 [87.00, 95.00]	0.704
RDW (%)	15.00 [13.80, 16.70]	15.10 [13.80, 16.70]	15.00 [13.90, 16.50]	0.324
Vital signs, median (IQR)				
Heartrate (min^{-1})	104.00 [90.00, 119.00]	104.00 [90.00, 119.00]	104.00 [91.00, 120.00]	0.071
Respiratory rate (min^{-1})	27.00 [23.50, 32.00]	27.00 [23.00, 32.00]	28.00 [24.00, 32.00]	0.033
Temperature (°C)	36.33 [35.61, 37.33]	36.33 [35.61, 37.33]	36.33 [35.61, 37.33]	0.555

explanation for specific patients using SHAP plots. The predictive value of a clinical factor for the XGBoost model increased with the average absolute SHAP value of each factor. Each factor was averaged to provide a homogeneous perspective, and the interpretation of SHAP was based on each individual patient (17). SHAP has two advantages: (1) it considers the effects of individual

TABLE 2 Research subject base information form (external validation).

	Total	Training cohorts	External validation cohorts	*p* value
N	6876	5,303	1,573	
Death = No/Yes (%)	6092/784 (88.6/11.4)	4690/613 (88.4/11.6)	1402/171 (89.1/10.9)	0.478
ICU stay time, days (median [IQR])	3.54 [2.04, 6.71]	3.38 [2.00, 6.33]	3.99 [2.28, 7.49]	<0.001
Age (median [IQR])	69.00 [58.00, 78.00]	69.00 [58.00, 79.00]	66.00 [55.00, 76.00]	<0.001
Gender = Male/Female (%)	4051/2825 (58.9/41.1)	3142/2161 (59.2/40.8)	909/664 (57.8/42.2)	0.315
Medical treatments, n (%)				
Sedatives = No/Yes	2093/4783 (30.4/69.6)	900/4403 (17.0/83.0)	1193/380 (75.8/24.2)	<0.001
Analgesic = No/Yes	1867/5009 (27.2/72.8)	492/4811 (9.3/90.7)	1375/198 (87.4/12.6)	<0.001
Antibiotic = No/Yes	2261/4615 (32.9/67.1)	689/4614 (13.0/87.0)	1572/1 (99.9/0.1)	<0.001
Vasopressin = No/Yes	6061/815 (88.1/11.9)	4634/669 (87.4/12.6)	1427/146 (90.7/9.3)	<0.001
Comorbidity, n (%)				
Myocardial infarct = No/Yes	5709/1167 (83.0/17.0)	4193/1110 (79.1/20.9)	1516/57 (96.4/3.6)	<0.001
Congestive heart failure = No/Yes	4855/2021 (70.6/29.4)	3397/1906 (64.1/35.9)	1458/115 (92.7/7.3)	<0.001
Peripheral vascular disease = No/Yes	6029/847 (87.7/12.3)	4468/835 (84.3/15.7)	1561/12 (99.2/0.8)	<0.001
Dementia = No/Yes	6710/166 (97.6/2.4)	5151/152 (97.1/2.9)	1559/14 (99.1/0.9)	<0.001
COPD = No/Yes	5370/1506 (78.1/21.9)	3814/1489 (71.9/28.1)	1556/17 (98.9/1.1)	<0.001
Liver disease = No/Yes	6245/631 (90.8/9.2)	4694/609 (88.5/11.5)	1551/22 (98.6/1.4)	<0.001
Diabetes = No/Yes	6368/508 (92.6/7.4)	4807/496 (90.6/9.4)	1561/12 (99.2/0.8)	<0.001
Renal disease = No/Yes	5360/1516 (78.0/22.0)	3935/1368 (74.2/25.8)	1425/148 (90.6/9.4)	<0.001
Severe score, median (IQR)				
SOFA	7.00 [5.00, 9.00]	7.00 [5.00, 9.00]	7.00 [5.00, 9.00]	0.043
GCS	13.00 [8.00, 14.00]	13.00 [8.00, 14.00]	11.00 [7.00, 14.00]	<0.001
APSIII	55.00 [40.00, 75.00]	54.00 [39.00, 75.00]	56.00 [42.00, 77.00]	<0.001
Laboratory tests, median (IQR)				
Lactate (mmol/L)	1.80 [1.20, 2.60]	1.80 [1.30, 2.60]	1.70 [1.10, 2.80]	0.011
Glucose (mg/dl)	127.75 [111.00, 145.29]	128.25 [113.00, 145.07]	124.00 [105.00, 146.00]	<0.001
Creatinine (mg/dl)	1.10 [0.79, 1.70]	1.05 [0.75, 1.65]	1.18 [0.82, 1.89]	<0.001
BUN (K/uL)	21.50 [14.50, 35.50]	21.00 [14.50, 35.00]	23.00 [14.00, 37.00]	0.068
Platelets (K/ul)	177.00 [126.50, 246.00]	180.00 [130.00, 249.00]	167.00 [117.00, 236.00]	<0.001
Potassium (K/ul)	4.20 [3.85, 4.60]	4.25 [3.90, 4.65]	4.10 [3.60, 4.51]	<0.001
Sodium (K/ul)	138.50 [136.00, 141.00]	138.50 [136.00, 141.00]	139.00 [136.00, 142.00]	0.029
Bicarbonate (meq/L)	23.00 [20.00, 25.50]	23.00 [20.50, 25.50]	23.00 [20.00, 25.00]	0.006
Calcium (mg/dl)	8.20 [7.70, 8.65]	8.20 [7.80, 8.70]	8.00 [7.40, 8.60]	<0.001
Chloride (mmol/L)	105.00 [101.00, 108.50]	105.00 [101.00, 108.00]	106.00 [102.00, 110.00]	<0.001
HCT (%)	10.90 [9.25, 14.40]	10.10 [8.85, 11.50]	30.60 [26.00, 35.90]	<0.001
PTT (s)	32.70 [28.10, 41.25]	32.60 [28.10, 41.90]	33.00 [28.20, 40.00]	0.582
WBC (K/uL)	12.10 [8.70, 16.35]	12.10 [8.85, 16.10]	12.00 [8.14, 17.10]	0.363
INR	1.30 [1.15, 1.55]	1.30 [1.15, 1.55]	1.30 [1.10, 1.60]	0.383
AG	13.50 [11.00, 16.00]	14.00 [12.00, 16.50]	11.00 [8.00, 15.00]	<0.001
MCH (pg)	30.00 [28.50, 31.20]	30.00 [28.60, 31.30]	29.80 [28.10, 31.10]	<0.001
MCHC (g/dL)	32.80 [31.70, 33.82]	32.80 [31.60, 33.90]	32.90 [31.90, 33.80]	0.227
MCV (fL)	91.00 [87.00, 95.00]	91.00 [87.00, 95.00]	90.00 [86.00, 94.60]	<0.001
RDW (%)	15.10 [13.90, 16.80]	15.10 [13.80, 16.70]	15.30 [14.00, 17.10]	<0.001
Vital signs, median (IQR)				
Heartrate (min^{-1})	101.00 [84.00, 117.00]	104.00 [90.00, 119.00]	92.00 [79.00, 109.00]	<0.001
Respiratory rate (min^{-1})	26.00 [18.00, 31.00]	27.00 [23.00, 32.00]	19.00 [15.00, 24.00]	<0.001
Temperature (°C)	36.39 [35.78, 37.28]	36.33 [35.61, 37.33]	36.70 [36.20, 37.10]	<0.001

factors and the synergy between factors, which can solve the multicollinearity problem, and (2) SHAP determines whether the influence is favorable (18).

In the current investigation, APSIII emerged as the most weighty contributor in the importance plot, underscoring its robust capacity to predict mortality in the ICU for individuals grappling with SAE. As

10.3389/fneur.2023.1290117

TABLE 3 Comparison of basic characteristics of the surviving and dead groups in the MIMIC-IV database.

Variable	Survival	Death	*p* value
Total	6709	867	
ICU stay time, days, median (IQR)	3.2 (1.9,5.8)	5.3 (2.6,10.6)	<0.001
Age, year, median (IQR)	69 (58,79)	74 (64,82)	<0.001
Gender, male, *n* (%)	4014 (59.8)	483 (55.7)	0.02
Medical treatments, n (%)			
Sedatives	5502 (82)	782 (90.2)	<0.001
Analgesic	6178 (92.1)	669 (77.2)	<0.001
Antibiotic	5790 (86.3)	798 (92)	<0.001
Vasopressin	564 (8.4)	370 (42.7)	<0.001
Comorbidity, n (%)			
Myocardial infarct	1396 (20.8)	183 (21.1)	0.838
Congestive heart failure	2362 (35.2)	377 (43.5)	<0.001
Peripheral vascular disease	1032 (15.4)	147 (17)	0.229
Dementia	192 (2.9)	31 (3.6)	0.242
COPD	1782 (26.6)	295 (34)	<0.001
Liver disease	710 (10.6)	195 (22.5)	<0.001
Diabetes	630 (9.4)	75 (8.7)	0.48
Renal disease	1661 (24.8)	288 (33.2)	<0.001
Severe score, median (IQR)			
SOFA	6 (4,9)	11 (8,14)	<0.001
GCS	13 (10,14)	7 (3,11)	<0.001
APSIII	51 (37,70)	90 (70,109)	<0.001
Laboratory tests, median (IQR)			
Lactate (mmol/L)	1.8 (1.3,2.5)	2.2 (1.4,3.7)	<0.001
Glucose (mg/dL)	128.5 (114.6,144.3)	128 (105.2,148.2)	0.02
Creatinine (mg/dL)	1 (0.8,1.6)	1.5 (0.9,2.5)	<0.001
BUN (K/uL)	20 (14,33)	33.5 (20,52.5)	<0.001
Platelets (K/uL)	179 (131,246.5)	176.5 (109.8,259.5)	0.042
Potassium (K/uL)	4.2 (3.9,4.6)	4.3 (3.9,4.9)	<0.001
Sodium (K/uL)	138.5 (136,141)	138.5 (135.5,141.5)	0.903
Bicarbonate (mEq/L)	23 (20.5,25.5)	21.5 (18,25)	<0.001
Calcium (mg/dL)	8.2 (7.8,8.7)	8.2 (7.6,8.7)	0.003
Chloride (mmol/L)	105.5 (101.5,108)	103.5 (99,107.5)	<0.001
HCT (%)	10.4 (1.9)	10.1 (2)	<0.001
PTT (s)	32 (27.9,40.2)	38.1 (30.5,53.5)	<0.001
WBC (K/uL)	11.9 (8.8,15.8)	13.2 (9,18.2)	<0.001
INR	1.3 (1.1,1.5)	1.5 (1.2,2)	<0.001
AG	13.5 (11.5,16)	16 (13.5,19.5)	<0.001
MCH (pg)	30 (28.6,31.3)	30 (28.5,31.5)	0.529
MCHC (g/dL)	32.8 (31.7,33.9)	32.3 (31,33.4)	<0.001
MCV (fL)	91 (87,95)	92 (88,97)	<0.001
RDW (%)	14.9 (13.7,16.4)	16.6 (15,18.3)	<0.001
Vital signs, median (IQR)			
Heartrate (min^{-1})	103 (90,118)	113 (97,130)	<0.001
Respiratory rate (min^{-1})	27 (23,32)	30 (25,35)	<0.001
Temperature (°C)	36.3 (35.6,37.3)	36.3 (35.6,37.3)	0.021

10.3389/fneur.2023.1290117

TABLE 4 Comparison of basic characteristics of the surviving and dead groups in the Eicu-CRD database.

Variable	Survival	Death	*p* value
Total	1402	171	
ICU stay time, days, median (IQR)	4 (2.3,7.4)	4.3 (2.3,8.2)	0.574
Age, year, median (IQR)	66 (54,76)	66 (57,77.5)	0.197
Gender, male, *n* (%)	821 (58.6)	88 (51.5)	0.076
Medical treatments, n (%)			
Sedatives	321 (22.9)	59 (34.5)	<0.001
Analgesic	162 (11.6)	36 (21.1)	<0.001
Antibiotic	1 (0.1)	0 (0)	0.727
Vasopressin	88 (6.3)	58 (33.9)	<0.001
Comorbidity, n (%)			
Myocardial infarct	48 (3.4)	9 (5.3)	0.224
Congestive heart failure	101 (7.2)	14 (8.2)	0.641
Peripheral vascular disease	11 (0.8)	1 (0.6)	0.777
Dementia	13 (0.9)	1 (0.6)	0.653
COPD	13 (0.9)	4 (2.3)	0.092
Liver disease	19 (1.4)	3 (1.8)	0.675
Diabetes	11 (0.8)	1 (0.6)	0.777
Renal disease	129 (9.2)	19 (11.1)	0.419
Severe score, median (IQR)			
SOFA	7 (5,9)	9 (7,12)	<0.001
GCS	11 (7,14)	9 (6,14)	<0.001
APSIII	55 (41,73)	77 (57,103)	<0.001
Laboratory tests, median (IQR)			
Lactate (mmol/L)	1.6 (1.1,2.6)	2.7 (1.5,5.1)	<0.001
Glucose (mg/dl)	124 (105,146)	123 (100,145)	0.531
Creatinine (mg/dl)	1.1 (0.8,1.8)	1.4 (1,2.3)	<0.001
BUN (K/uL)	22 (14,35)	30 (20,48)	<0.001
Platelets (K/ul)	166 (118,234)	171 (91.5,243)	0.416
Potassium (K/ul)	4.1 (3.6,4.5)	4.1 (3.6,4.8)	0.226
Sodium (K/ul)	139 (136,142)	138 (135,142)	0.192
Bicarbonate (meq/L)	23 (20,25)	21 (17,24)	<0.001
Calcium (mg/dl)	8 (7.4,8.6)	8 (7.4,8.4)	0.102
Chloride (mmol/L)	106 (102,110)	105 (100,109)	0.096
HCT (%)	30.7 (26,36)	29.8 (25.8,34.7)	0.188
PTT (s)	32.9 (28,39.4)	35.8 (30.7,45.1)	<0.001
WBC (K/uL)	11.9 (8.2,17)	13 (8.1,18.9)	0.231
INR	1.3 (1.1,1.6)	1.5 (1.2,2.1)	<0.001
AG	11 (8,14)	14 (10,16)	<0.001
MCH (pg)	29.8 (28.1,31.1)	29.4 (27.9,31.1)	0.331
MCHC (g/dL)	32.9 (32,33.8)	32.5 (31.4,33.3)	<0.001
MCV (fL)	90 (86.2,94.4)	91.2 (85.6,96)	0.319
RDW (%)	15.1 (14,17)	16.4 (14.9,18.2)	<0.001
Vital signs, median (IQR)			
Heartrate (min^{-1})	92 (79,108.8)	93 (80.5,110)	0.231
Respiratory rate (min^{-1})	19 (15,24)	20 (16,25)	0.014
Temperature (°C)	36.7 (36.3,37.2)	36.5 (36,37)	<0.001

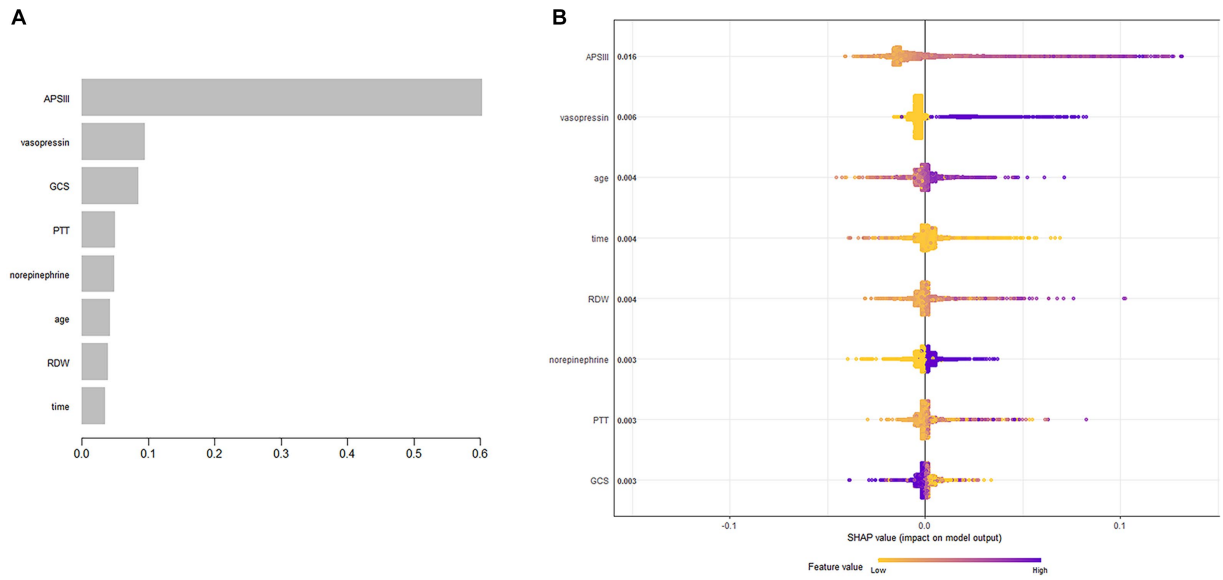


FIGURE 2
Predictor variables selection. **(A)** Importance of the predictor variables selected by XGBoost. **(B)** The SHAP summary plot.

an integral facet of the APACHE system, APSIII aptly showcases its aptitude in prognosticating the mortality rates of patients grappling with severe sepsis and septic shock (19). In a large number of studies it has been found that survivors have significantly lower APACHE III scores than deceased patients and that higher scores (OR 1.11,95% CI 1.05–1.18, $p = 0.001$) are associated with increased in-hospital all-cause mortality in patients with severe sepsis (20, 21). Inflammatory response, immunosuppression, and multiple organ dysfunction syndrome may be responsible for the high scores in patients with SAE (1, 22). A predominant clinical characteristic of SAE is the alteration in the level of consciousness. In milder cases, there is a reduction in attention and alertness, accompanied by symptoms like anxiety and delirium. In severe instances, it may lead to stupor or coma. Long-term cognitive impairments encompass deficits in memory, attention, verbal fluency, and executive functions, significantly impacting the quality of life for survivors. In a study concentrated on discerning initial and potentially amendable factors of SAE upon admission to the ICU, it was determined that even slight alterations in cognitive function, as defined by a GCS score of 13–14, were autonomously correlated with mortality at the point of ICU admission (10, 23). Furthermore, our findings confirm the independent role of the GCS score as a risk factor for ICU mortality in SAE patients. This reinforces the utility of the GCS score and APSIII in gauging the severity and prognosis of individuals afflicted with SAE.

Norepinephrine and vasopressin are now commonly used in clinics as vasoactive drugs. In the Surviving Sepsis Campaign guidelines, norepinephrine is recommended as the vasopressor for sepsis treatment (24), and often vasopressin is used as an adjunct to sepsis. Maheshwari et al. found that the significant blood pressure response to VAS was substantially linked to reduced survival probability in patients with septic shock (25). However, there are only treatment guidelines for sepsis and septic shock (15, 17), with a lack of specific treatment guidelines for SAE (24, 26). To realistically assess ICU mortality in SAE, clinicians should be aware of other treatment options for SAE in order to improve the corresponding survival assessment system. Currently, no specific therapeutic interventions are tailored for SAE. Treatment protocols are established on the comprehensive management of sepsis, with a predominant focus on symptoms associated with cerebral maladies, while endeavoring to minimize detriment to the central nervous system. Early-stage resuscitation is acknowledged as a pivotal therapeutic strategy for sepsis, and the administration of vasoactive agents correlated with normal arterial pressure subsequent to initial fluid therapy can mitigate the severity of sepsis (27). Furthermore, glucocorticoids, alternative markers, and modulators of the neuroimmune axis have been under consideration for addressing sepsis-induced cognitive impairments (28, 29). Indoleamine 2,3-dioxygenase, impacting the inflammatory cascade, is identified as a potential therapeutic target for central nervous system disorders, fostering cognitive enhancement in sepsis patients (30).

Most of the patients with SAE in this study stayed in the ICU for less than 1 week, and the length of ICU stay had an overall negative effect on the outcome. Related studies have found length of ICU stay to be related to disease severity (31), which has important implications for the wise use of scarce medical resources (32). Elderly patients with SAE admitted to the ICU mostly died earlier than did younger patients with SAE, which is supported by the findings of Martin et al. (33). The introduction of comorbidities harm immune function as age progresses, which causes patients with critical illness to deteriorate more rapidly. Geriatric patients may be more vulnerable to CNS issues, particularly if hypertension, diabetes mellitus, or acute renal injury is the underlying illness (10, 34, 35). Older hospitalized people need more-specialized nursing or rehabilitation care. These findings offer guidance on how to allocate healthcare resources for patients with SAE and offer suggestions for future research projects and patient interventions.

RDW may be therapeutically valuable for predicting the future course and prognosis of various disorders, including stroke, atrial fibrillation (36), COPD (37), community-acquired pneumonia, and sepsis (38). Currently, several studies have indicated that RDW

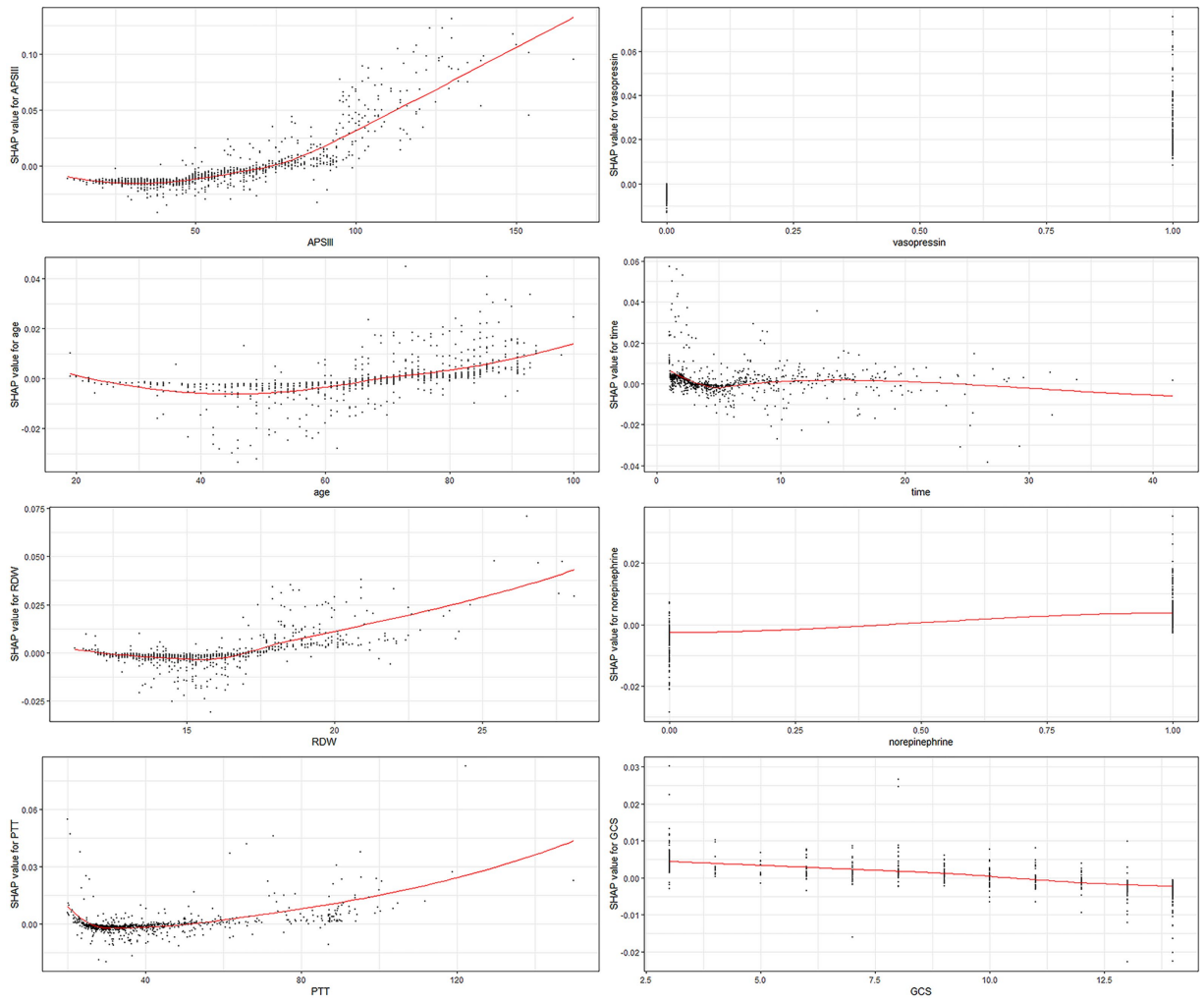


FIGURE 3
SHAP dependency plot of the XGboost model. The SHAP dependence plot shows how a single feature affects the output of the XGBoost prediction model. SHAP values for specific features exceed zero, representing an increased risk of death. RDW, red blood cell distribution width; PTT, partial thromboplastin time.

possesses not only diagnostic significance but also serves as a prognostic factor for sepsis, signifying systemic dysfunction and immune response dysregulation. Víctor Moreno-Torres, MD, et al. have proposed that RDW enhances the discriminative capacity of SOFA, LODS, APACHE-II, and SAPS-II, rendering it a potential parameter within these prognostic scoring systems (39). A study by Sadaka et al., which included 279 patients with septic shock, suggested that elevated RDW at admission was related to death in the ICU in both adults and neonates (40). This finding was also demonstrated in another study, which found that the addition of RDW to the ICU scoring system improved its mortality predictions (41, 42). The present study indicated there was a higher risk of dying from SAE in the ICU when RDW levels were high. Elevated RDW reflects a severe dysregulation of red blood cell homeostasis, which may be an important prognostic factor for SAE. The mechanistic relationship between RDW and the ICU mortality rate in SAE remains obscure. However, research suggests that oxidative stress may contribute to the detrimental impact of RDW on the prognosis of SAE, as oxidative stress levels exhibit a positive correlation with RDW (43). Apart from this, the inflammatory response in septic patients shortens red blood cell lifespan, impairs red blood cell maturation, resulting in premature release, and thus elevating RDW (44, 45). Furthermore, proinflammatory cytokines inhibit erythropoietin-induced red blood cell proliferation and maturation, also leading to an increase in RDW (46). This may represent another rationale for the association between RDW and ICU mortality.

One strength of this study was the external validation of the SAE mortality risk model using the eICU-CRD database, which confirmed its efficacy. SHAP allows visualization of XGBoost models, and its sound visual interpretation greatly increases the confidence that clinicians have in the application of machine learning. However, there were some limitations to this study. First, only data from the US were utilized to construct and validate the model, which might reduce its applicability to other regions of the world. Furthermore, in retrospective studies, it is inevitable to relinquish certain variables with a substantial amount of missing values. Various unmeasured

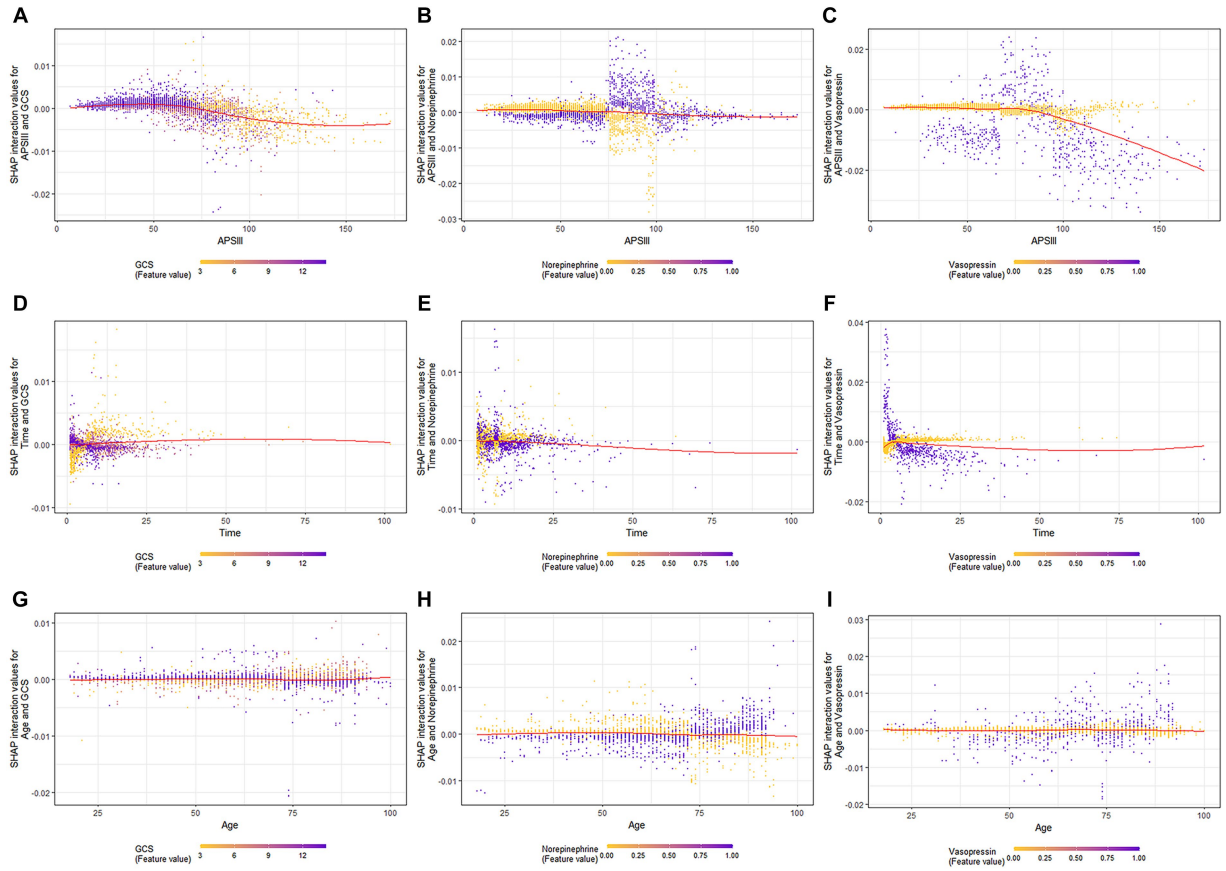


FIGURE 4
SHAP interaction plot of the eight most essential features for SAE assessment. SHAP, SHapley Additive explanation; XGBoost, eXtreme Gradient Boosting; SAE, sepsis-associated encephalopathy.

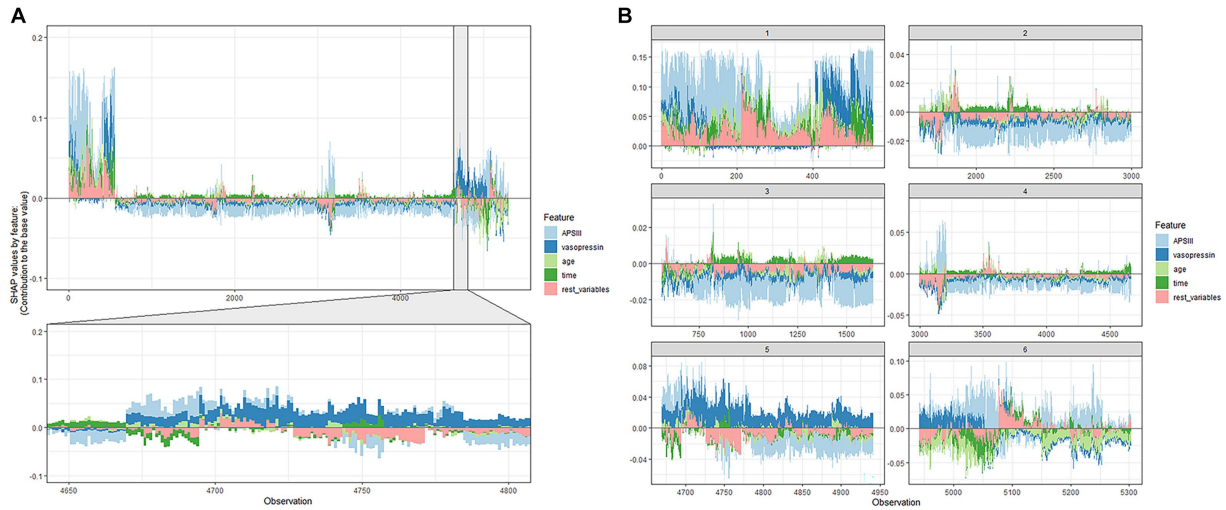


FIGURE 5
SHAP force plot of the XGboost model. **(A)** Influence plot of macroscopic features for all samples. **(B)** Influence plot of macroscopic features for a random portion of the samples. A positive Shap value represents a positive gain area and a negative Shap value represents a negative gain area.

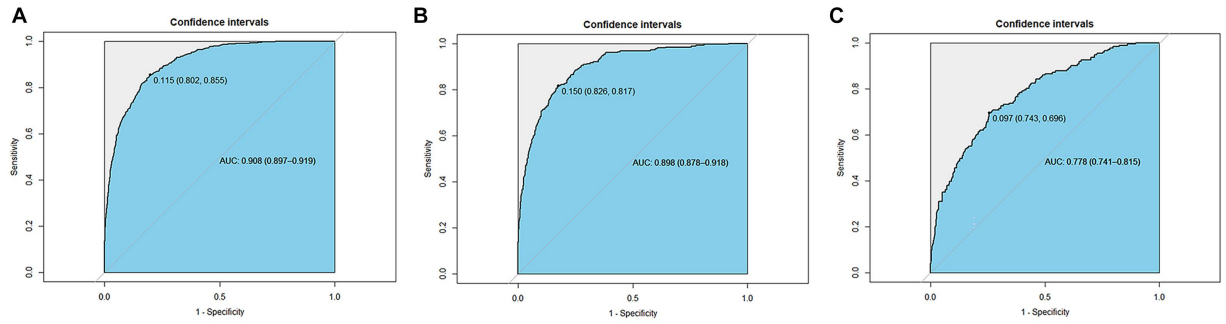


FIGURE 6
Receiver operating characteristic curves of the XGBoost model. **(A)** The test set (AUC=0.908). **(B)** The internal validation set (AUC=0.898). **(C)** The external validation set (AUC=0.778).

confounding variables such as race and treatment modalities, along with inflammation-related data, could potentially influence the mortality risk of SAE patients. Therefore, given the constraints of the MIMIC-IV database and the eICU-CRD database, it is plausible that the XGBoost model may have omitted certain pivotal factors. Finally, the divergence in the origins of these two patient cohorts, hailing from distinct databases, has led to disparate study timelines between the training and external validation cohorts. This temporal incongruity might be a pivotal element contributing to the attenuation of this model's efficacy within the eICU-CRD database.

Conclusion

This study validated the efficacy of machine-learning-based XGBoost for early outcome predictions for patients with SAE. The SHAP method improves the readability of XGBoost models and aids doctors in comprehending the logic behind findings obtained from such models.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The article does not belong to the scope of the Ethics Committee Review and does not need to be reviewed according to the current ethical standards. Both MIMIC-IV and ICU-CRD are public databases and are public resources. The authors have passed the application of both MIMIC-IV and ICU-CRD databases, completed the CITI examination and account registration, and have the permission to use them.

Author contributions

JG: Validation, Writing – original draft, Writing – review & editing. HC: Writing – original draft, Writing – review & editing. ZW: Methodology, Writing – review & editing. MQ: Validation, Writing – review & editing. JiL: Conceptualization, Writing – review & editing. JuL: Funding acquisition, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization (2021B1212040007), the Clinical Frontier Technology Program of the First Affiliated Hospital of Jinan University, China (No. JNU1AF-CFTP-2022-a01235), and the Science and Technology Projects in Guangzhou, China (Nos. 202201020054 and 2023A03J1032).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Ren C, Yao R-Q, Zhang H, Feng Y-W, Yao Y-M. Sepsis-associated encephalopathy: a vicious cycle of immunosuppression. *J Neuroinflammation*. (2020) 17:14. doi: 10.1186/s12974-020-1701-3

2. Rabuel C, Mebazaa A. Septic shock: a heart story since the 1960s. *Intensive Care Med*. (2006) 32:799–807. doi: 10.1007/s00134-006-0142-5

3. Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, et al. Incidence and outcomes of acute lung injury. *N Engl J Med*. (2005) 353:1685–93. doi: 10.1056/NEJMoa050333

4. Gofton TE, Young GB. Sepsis-associated encephalopathy. *Nat Rev Neurol*. (2012) 8:557–66. doi: 10.1038/nrneurol.2012.183

5. Sonneville R, Verdonk F, Rauturier C, Klein IF, Wolff M, Annane D, et al. Understanding brain dysfunction in sepsis. *Ann Intensive Care*. (2013) 3:15. doi: 10.1186/2110-5820-3-15

6. Handa O, Stephen J, Cepinskas G. Role of endothelial nitric oxide synthase-derived nitric oxide in activation and dysfunction of cerebrovascular endothelial cells during early onsets of sepsis. *Am J Physiol Heart Circ Physiol*. (2008) 295:H1712–9. doi: 10.1152/ajpheart.00476.2008

7. Gao Q, Hernandes MS. Sepsis-associated encephalopathy and blood-brain barrier dysfunction. *Inflammation*. (2021) 44:2143–50. doi: 10.1007/s10753-021-01501-3

8. Chen J, Shi X, Diao M, Jin G, Zhu Y, Hu W, et al. A retrospective study of sepsis-associated encephalopathy: epidemiology, clinical features and adverse outcomes. *BMC Emerg Med*. (2020) 20:77. doi: 10.1186/s12873-020-00374-3

9. Chung H-Y, Wickel J, Brunkhorst FM, Geis C. Sepsis-associated encephalopathy: from delirium to dementia? *J Clin Med*. (2020) 9:E703. doi: 10.3390/jcm9030703

10. Sonneville R, de Montmollin E, Poujade J, Garrouste-Orgeas M, Souweine B, Darmon M, et al. Potentially modifiable factors contributing to sepsis-associated encephalopathy. *Intensive Care Med*. (2017) 43:1075–84. doi: 10.1007/s00134-017-4807-z

11. Johnson AEW, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. (2016) 3:160035. doi: 10.1038/sdata.2016.35

12. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. (2015) 350:g7594. doi: 10.1136/bmj.g7594

13. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU collaborative research database, a freely available multi-center database for critical care research. *Sci Data*. (2018) 5:180178. doi: 10.1038/sdata.2018.178

14. Liu J, Wu J, Liu S, Li M, Hu K, Li K. Predicting mortality of patients with acute kidney injury in the ICU using XGBoost model. *PLoS One*. (2021) 16:e0246306. doi: 10.1371/journal.pone.0246306

15. Zheng H, Yuan J, Chen L. Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. *Energies*. (2017) 10:1168. doi: 10.3390/en10081168

16. Chen H, Lundberg S, Lee S-I. Explaining models by propagating Shapley values of local components. *Explain AI Healthc Med*. (2021) 914:261–70. doi: 10.1007/978-3-030-53352-6_24

17. Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H, et al. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. *Ann Transl Med*. (2018) 6:216. doi: 10.21037/atm.2018.05.32

18. Zou Y, Shi Y, Sun F, Liu J, Guo Y, Zhang H, et al. Extreme gradient boosting model to assess risk of central cervical lymph node metastasis in patients with papillary thyroid carcinoma: individual prediction using SHapley additive exPlanations. *Comput Methods Prog Biomed*. (2022) 225:107038. doi: 10.1016/j.cmpb.2022.107038

19. Arabi Y, Al Shirawi N, Memish Z, Venkatesh S, Al-Shimemeri A. Assessment of six mortality prediction models in patients admitted with severe sepsis and septic shock to the intensive care unit: a prospective cohort study. *Crit Care*. (2003) 7:R116–22. doi: 10.1186/cc2373

20. Rogy MA, Oldenburg HSA, Coyle S, Trousdale R, Moldawer LL, Lowry SF. Correlation between acute physiology and chronic health evaluation (APACHE) III score and immunological parameters in critically ill patients with sepsis. *Br J Surg*. (2005) 83:396–400. doi: 10.1002/bjs.1800830333

21. Dobesh PP, Klepser DG, McGuire TR, Morgan CW, Olsen KM. Reduction in mortality associated with statin therapy in patients with severe Sepsis. *Pharmacotherapy*. (2009) 29:621–30. doi: 10.1592/phco.29.6.621

22. Kaml GJ, Davis KA. Surgical critical care for the patient with sepsis and multiple organ dysfunction. *Anesthesiol Clin*. (2016) 34:681–96. doi: 10.1016/j.anclin.2016.06.005

23. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe Sepsis. *JAMA*. (2010) 304:1787–94. doi: 10.1001/jama.2010.1553

24. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. (2021) 47:1181–247. doi: 10.1007/s00134-021-06506-y

25. Maheshwari K, Nathanson BH, Munson SH, Khangulov V, Stevens M, Badani H, et al. The relationship between ICU hypotension and in-hospital mortality and morbidity in septic patients. *Intensive Care Med*. (2018) 44:857–67. doi: 10.1007/s00134-018-5218-5

26. Russell JA, Gordon AC, Williams MD, Boyd JH, Walley KR, Kissoon N. Vasopressor therapy in the intensive care unit. *Semin Respir Crit Care Med*. (2021) 42:059–77. doi: 10.1055/s-0040-1710320

27. Pan S, Lv Z, Wang R, Shu H, Yuan S, Yu Y, et al. Sepsis-induced brain dysfunction: pathogenesis, diagnosis, and treatment. *Oxidative Med Cell Longev*. (2022) 2022:1328729–13. doi: 10.1155/2022/1328729

28. Ma Y, Matsuwaki T, Yamanouchi K, Nishihara M. Glucocorticoids suppress the protective effect of Cyclooxygenase-2-related signaling on hippocampal neurogenesis under acute immune stress. *Mol Neurobiol*. (2017) 54:1953–66. doi: 10.1007/s12035-016-9766-9

29. Ağaç D, Estrada LD, Maples R, Hooper LV, Farrar JD. The β2-adrenergic receptor controls inflammation by driving rapid IL-10 secretion. *Brain Behav Immun*. (2018) 74:176–85. doi: 10.1016/j.bbi.2018.09.004

30. Comim CM, Freiberger V, Ventura L, Mina F, Ferreira GK, Michels M, et al. Inhibition of indoleamine 2,3-dioxygenase 1/2 prevented cognitive impairment and energetic metabolism changes in the hippocampus of adult rats subjected to polymicrobial sepsis. *J Neuroimmunol*. (2017) 305:167–71. doi: 10.1016/j.jneuroim.2017.02.001

31. Zampieri FG, Ladeira JP, Park M, Haib D, Pastore CL, Santoro CM, et al. Admission factors associated with prolonged (>14 days) intensive care unit stay. *J Crit Care*. (2014) 29:60–5. doi: 10.1016/j.jcrc.2013.09.030

32. Takekawa D, Endo H, Hashiba E, Hirota K. Predict models for prolonged ICU stay using APACHE II, APACHE III and SAPS II scores: a Japanese multicenter retrospective cohort study. *PLoS One*. (2022) 17:e0269737. doi: 10.1371/journal.pone.0269737

33. Martin GS, Mannino DM, Moss M. The effect of age on the development and outcome of adult sepsis. *Crit Care Med*. (2006) 34:15–21. doi: 10.1097/01.ccm.0000194535.82812.ba

34. Peters R, Collerton J, Granic A, Davies K, Kirkwood T, Jagger C. Antihypertensive drug use and risk of cognitive decline in the very old: an observational study – the Newcastle 85+ study. *J Hypertens*. (2015) 33:2156–64. doi: 10.1097/HJH.0000000000000653

35. Sonneville R, Vanhorebeek I, den Hertog HM, Chrétien F, Annane D, Sharshar T, et al. Critical illness-induced dysglycemia and the brain. *Intensive Care Med*. (2015) 41:192–202. doi: 10.1007/s00134-014-3577-0

36. Fava C, Cattazzo F, Hu Z-D, Lippi G, Montagnana M. The role of red blood cell distribution width (RDW) in cardiovascular risk assessment: useful or hype? *Ann Transl Med*. (2019) 7:581. doi: 10.21037/atm.2019.09.58

37. Salvagno GL, Sanchis-Gomar F, Picanza A, Lippi G. Red blood cell distribution width: a simple parameter with multiple clinical applications. *Crit Rev Clin Lab Sci*. (2015) 52:86–105. doi: 10.3109/10408363.2014.992064

38. Hu Z-D, Lippi G, Montagnana M. Diagnostic and prognostic value of red blood cell distribution width in sepsis: a narrative review. *Clin Biochem*. (2020) 77:1–6. doi: 10.1016/j.clinbiochem.2020.01.001

39. Moreno-Torres V, Royuela A, Múñez-Rubio E, Gutierrez-Rojas Á, Mills-Sánchez P, Ortega A, et al. Red blood cell distribution width as prognostic factor in sepsis: a new use for a classical parameter. *J Crit Care*. (2022) 71:154069. doi: 10.1016/j.jcrc.2022.154069

40. Sadaka F, O'Brien J, Prakash S. Red cell distribution width and outcome in patients with septic shock. *J Intensive Care Med*. (2013) 28:307–13. doi: 10.1177/0885066612452838

41. Meynaar IA, Knook AHM, Coolen S, Le H, Bos MMEM, van der Dijs F, et al. Red cell distribution width as predictor for mortality in critically ill patients. *Neth J Med*. (2013) 71:488–93.

42. Loveday S, Sinclair L, Badrick T. Does the addition of RDW improve current ICU scoring systems? *Clin Biochem*. (2015) 48:569–74. doi: 10.1016/j.clinbiochem.2015.04.002

43. Patel KV, Semba RD, Ferrucci L, Newman AB, Fried LP, Wallace RB, et al. Red cell distribution width and mortality in older adults: a meta-analysis. *J Gerontol A Biol Sci Med Sci*. (2010) 65A:258–65. doi: 10.1093/gerona/glp163

44. Weiss G, Goodnough LT. Anemia of chronic disease. *N Engl J Med*. (2005) 352:1011–23. doi: 10.1056/NEJMra041809

45. Pierce CN, Larson DF. Inflammatory cytokine inhibition of erythropoiesis in patients implanted with a mechanical circulatory assist device. *Perfusion*. (2005) 20:83–90. doi: 10.1191/0267659105pf793oa

46. Krishna V, Pillai G, Velickakathu SS. Red cell distribution width as a predictor of mortality in patients with Sepsis. *Cureus*. (2021) 13:e12912. doi: 10.7759/cureus.12912

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 18 December 2023
DOI 10.3389/fneur.2023.1282833

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Eduardo C. Gonzalez-Toledo,
Louisiana State University, United States
Thomas Welton,
National Neuroscience Institute (NNI), Singapore

*CORRESPONDENCE
Jose G. Tamez-Pena
jose.tamezpena@tec.mx

RECEIVED 24 August 2023
ACCEPTED 01 December 2023
PUBLISHED 18 December 2023

CITATION
Meyers SP, Hirad A, Gonzalez P, Bazarian JJ, Mirabelli MH, Rizzone KH, Ma HM, Rosella P, Totterman S, Schreyer E and Tamez-Pena JG (2023) Clinical performance of a multiparametric MRI-based post concussive syndrome index.
Front. Neurol. 14:1282833.
doi: 10.3389/fneur.2023.1282833

COPYRIGHT
© 2023 Meyers, Hirad, Gonzalez, Bazarian, Mirabelli, Rizzone, Ma, Rosella, Totterman, Schreyer and Tamez-Pena. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Clinical performance of a multiparametric MRI-based post concussive syndrome index

Steven P. Meyers[1], Adnan Hirad[2], Patricia Gonzalez[3], Jeffrey J. Bazarian[4], Mark H. Mirabelli[5], Katherine H. Rizzone[5], Heather M. Ma[6], Peter Rosella[1], Saara Totterman[3], Edward Schreyer[3] and Jose G. Tamez-Pena[7]*

[1]Department of Imaging Sciences, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [2]Department of Vascular Surgery, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [3]Qmetrics Technologies, Pittsford, NY, United States, [4]Departments of Emergency Medicine, Neurology, Neurosurgery, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [5]Department of Orthopedics, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [6]Department of Physical Medicine and Rehabilitation, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [7]School of Medicine and Health Sciences, Tecnologico de Monterey, Monterrey, Mexico

Introduction: Diffusion Tensor Imaging (DTI) has revealed measurable changes in the brains of patients with persistent post-concussive syndrome (PCS). Because of inconsistent results in univariate DTI metrics among patients with mild traumatic brain injury (mTBI), there is currently no single objective and reliable MRI index for clinical decision-making in patients with PCS.

Purpose: This study aimed to evaluate the performance of a newly developed PCS Index (PCSI) derived from machine learning of multiparametric magnetic resonance imaging (MRI) data to classify and differentiate subjects with mTBI and PCS history from those without a history of mTBI.

Materials and methods: Data were retrospectively extracted from 139 patients aged between 18 and 60 years with PCS who underwent MRI examinations at 2 weeks to 1-year post-mTBI, as well as from 336 subjects without a history of head trauma. The performance of the PCS Index was assessed by comparing 69 patients with a clinical diagnosis of PCS with 264 control subjects. The PCSI values for patients with PCS were compared based on the mechanism of injury, time interval from injury to MRI examination, sex, history of prior concussion, loss of consciousness, and reported symptoms.

Results: Injured patients had a mean PCSI value of 0.57, compared to the control group, which had a mean PCSI value of 0.12 ($p = 8.42e\text{-}23$) with accuracy of 88%, sensitivity of 64%, and specificity of 95%, respectively. No statistically significant differences were found in the PCSI values when comparing the mechanism of injury, sex, or loss of consciousness.

Conclusion: The PCSI for individuals aged between 18 and 60 years was able to accurately identify patients with post-concussive injuries from 2 weeks to 1-year post-mTBI and differentiate them from the controls. The results of this study suggest that multiparametric MRI-based PCSI has great potential as an objective clinical tool to support the diagnosis, treatment, and follow-up care of patients with post-concussive syndrome. Further research is required to investigate the replicability of this method using other types of clinical MRI scanners.

KEYWORDS

post-concussive syndrome, MTBI, MRI, radiomics, diffusion imaging

Introduction

Traumatic brain injuries (TBIs) are evaluated based on clinical symptoms, neurological impairments, and imaging findings (1–5). The majority (80–90%) of TBIs are mild, characterized by Glasgow Coma Scale scores of 13–15 (5–7). Mild TBI (mTBI) and concussion are often used interchangeably, with sports-related concussions being a subtype (6, 8). It is estimated that 1.4–3.8 million concussions occur annually in the US and are caused by sports, recreational activities, falls, assaults, and motor vehicle accidents (1, 2, 4, 5).

Symptoms associated with concussions include headaches, amnesia, dizziness, fatigue, drowsiness, sleep disturbance, irritability, blurred vision, nausea, hypersensitivity to light and noise, emotional lability, anxiety, depression, deficits in attention, concentration, memory, executive function, balance problems, and/or loss of consciousness for less than 30 min (5, 7–14). Loss of consciousness occurs in 10–20% of concussions but is not required for diagnosis (15–17). Most symptoms resolve within 14 days in adults and 4–6 weeks in children (4, 6, 7, 16–18). However, 15–30% of mTBI patients may experience post-concussion symptoms, referred to as persistent post-concussive syndrome (PCS), for several months or longer (4, 6, 7, 17–20). In the Prospective TRACK-TBI study, 33% of patients remained functionally impaired 3 months after injury, and 22.4% were not at full functional status 1 year after injury (4), further confirming that some patients with mTBI may experience long-term disability (7). Schneider et al. reported a poor 1-year cognitive outcome in 13.5% of patients with mTBI (21).

In most post-acute patients with a history of mTBI and PCS, there is no radiological evidence of brain injury using computed tomography (CT) or conventional MRI techniques (4, 5, 7, 18, 22). Recently, an advanced MRI technique, Diffusion Tensor Imaging (DTI), has been used to evaluate mild traumatic injuries in the acute, subacute, and delayed phases (5, 7, 18). However, literature reviews have found variable, inconsistent, or negative findings in the diffusion metrics between patients with PCS and controls (5, 6, 18, 22). Additionally, in patients with PCS or post-injury behavioral changes (18), diffusion imaging abnormalities were inconsistent. The inconsistencies in group differences in the locations of DTI-related white matter abnormalities have been proposed to be related to the heterogeneous nature and symptoms of mTBI, different mechanisms of injury, variable locations and phases of injury, differences in DTI protocols, and/or the limited numbers of control and subject populations (5, 18, 23). Previous studies using DTI to evaluate PCS at the group level have focused on univariate analyses of diffusion metrics such as FA, ADC, MD, RD, and AD (24). Although microstructural changes in DTI metrics have been demonstrated at the group level in patients with PCS, the complexity of advanced DTI post-analysis limits its application of these changes at the individual level in clinical settings for diagnosis, treatment, and follow-up care (18, 25).

Despite advances in measuring changes in the brain related to traumatic brain injury, there is currently no objective and reliable MRI assessment to guide clinical decision-making for individual patients with mTBI and PCS (18, 24, 25). The lack of objective data on mTBI in patients leads to challenges in the diagnosis, prognosis, and treatment of patients with a history of mTBI and PCS, based on subjective symptom reports and clinical examinations (25). To address this issue, machine learning (ML) approaches have been suggested (26–28). ML has also been applied to MRI data from patients with traumatic brain injuries. For example, Mitra et al. used a ML technique to classify patients with a history of mild, moderate, or severe TBI, based on altered structural connectivity patterns within intra-and interhemispheric white matter pathways secondary to trauma (28). Additionally, Goswami et al. reported that for retired football players with a history of multiple concussions, ML of mean and radial diffusion data showed alterations involving the uncinate fasciculus, which is associated with behavioral regulation (29). Vergara et al. reported that data from resting-state fMRI used to assess network connectivity was more accurate than DTI in detecting mild traumatic brain injury at the group level (30). Luo et al. reported that a support vector machine algorithm of multiparametric fMRI data in patients with mild traumatic brain injury could improve the classification performance of mTBI compared to normal controls by using the brain regions associated with emotion and cognition (27). Lui et al. (24) reported that an algorithm developed from multi-feature analysis of data from diffusion-weighted imaging, fMRI, and volumetrics may aid in the classification of patients with mTBI compared to controls (24). Abdelrahman et al. reported that combining multiple DTI metrics improved the accuracy of identifying patients with chronic moderate brain injury, with a mean time since injury of 9 years, compared with controls (31).

ML has also been widely applied to multiparametric clinical MRI data in oncology and other medical conditions (32, 33). The purpose of using ML has been to develop clinical tools that support diagnoses, predict prognoses, and predict responses to treatment for different diseases and medical conditions (33). In our previous work, we developed a PCS index (PCSI) for patients with a history of mTBI, using a feature selection process applied to multiparametric structural and diffusion MRI data. The PCSI combines complex radiomic information from the MP RAGE series, fractional anisotropy (FA), and apparent diffusion coefficient (ADC) series (26). This method enabled the detection of post-concussive imaging changes in all series, even when no apparent findings were evident on clinical MRI exams.

The objective of this study was to further evaluate the performance of a multiparametric MRI-based PCSI to detect post-concussive injury in a real-world clinical environment. A significant limitation of the previous study was the control population, which consisted of a relatively small number of young, healthy control athletes. Here, we expand the application of the PSCI to a significantly larger number of routine clinical patients diagnosed with mTBI and patients referred for neuroimaging without history of prior brain injury to explore the associations of the PCSI with sex, mechanism of injury, elapsed time from injury, prior concussion history and clinical symptoms in patients aged 18–60.

Materials and methods

Participants

The Research Subjects Review Board of the University of Rochester approved this retrospective study. Figure 1 shows the

Abbreviations: AI, artificial intelligence; ML, machine learning; mTBI, mild traumatic brain injury; SRC, sports related concussion; PCS, post-concussion syndrome; DTI, diffusion weighted imaging; FA, fractional anisotropy mapping from DTI sequences; ADC, apparent diffusion coefficient from DTI sequences.

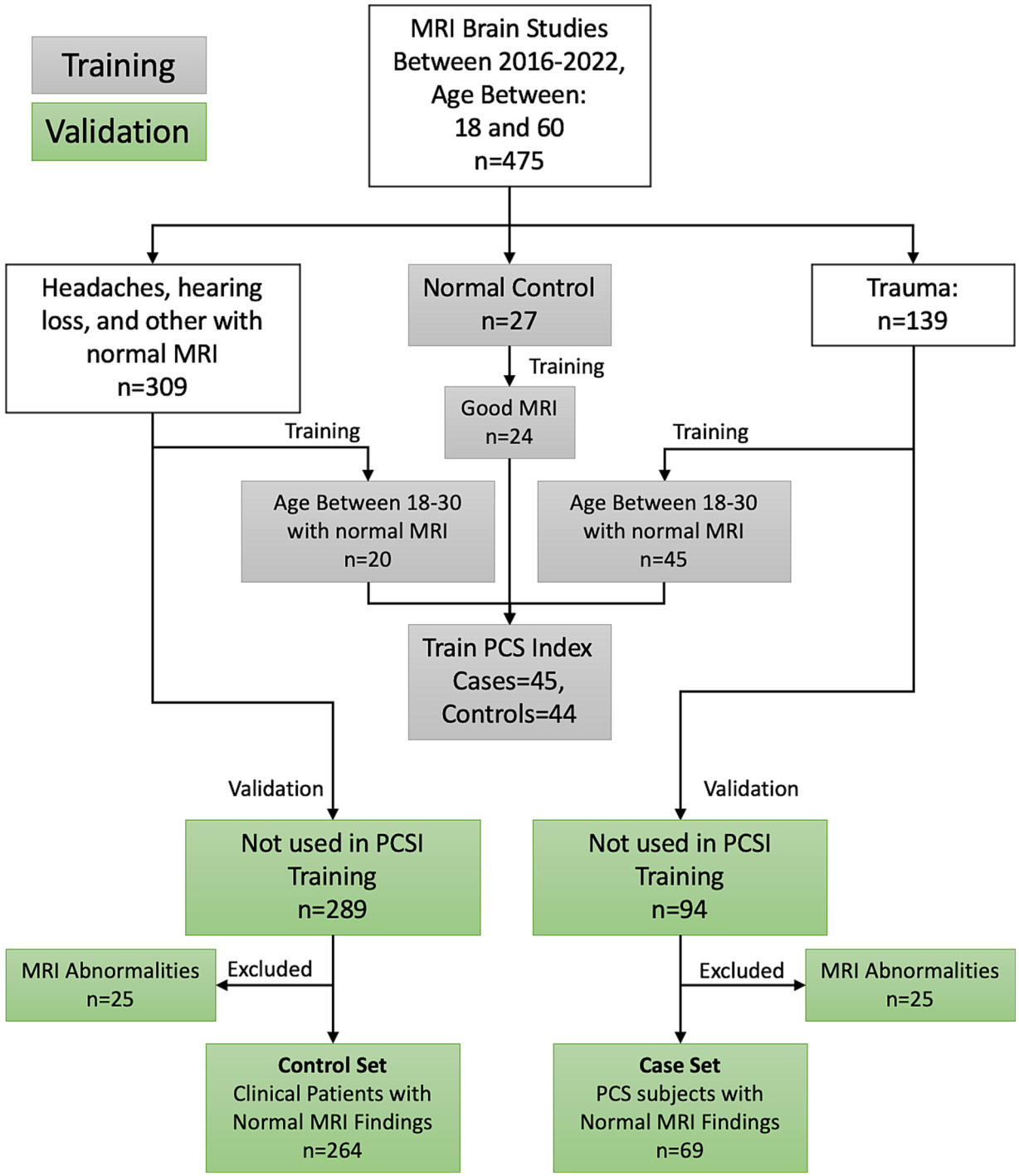


FIGURE 1
Inclusion–exclusion criteria for the study participants. This paper presents the performance of the PCSI on the validation participants.

inclusion and exclusion criteria for this retrospective review of the medical records of patients referred to an outpatient MRI center from 2016 to 2022. MRI was performed for 309 uninjured patients ranging in age from 18 to 60 years who were referred for MRI because of non-trauma related subjective complaints of headaches ($n = 88$), hearing loss ($n = 132$), or other complaints ($N = 42$), and who had normal MRI examinations. The uninjured patients also included 27 student athletes recruited as part of the prior study (26). A retrospective review was also performed of the medical records from 2016 to 2022 for 139 patients ranging in age from 18 to 60 years who were referred for MRI at the same facility by Sports Medicine Physicians, Physical Medicine, and Rehabilitation Physicians, or Neurologists with the diagnosis of concussion based on clinical history and combination of symptoms (4, 6–8, 10–18). The inclusion criteria included a clinical diagnosis of concussion with persistent post-concussion symptoms (PCS) and MRI performed at least 2 weeks

and no later than 12 months after mTBI, excluding those subjects with dental braces, prior brain surgery, ventricular shunts, intracranial hemorrhage, intra-axial MRI signal abnormalities, skull fractures, or standard contraindications for MR. Thirty eight patients received MR exams within 90 days from injury; 17 between 90 and 180 days from injury, 7 between 180 and 270 days from injury, with the remaining 7 between 270 days and 1 year from injury. Twenty five injured and 25 uninjured subjects were excluded due to issues with the MRI exams, such as missing series required for analysis, or excessive artifact. Patient age and sex, number of previous concussions, absence of loss of consciousness (LOC) at the most recent concussion, the time between injury and MRI, and persistent signs and symptoms were extracted from the electronic medical records. All patient identifiers were removed prior to the analysis.

Image acquisition

All MRI exams for the concussed patients and 27 student athlete controls were performed between 2016, and 2022, on one of two 3 T Siemens Skyra MRI Scanners using a 20-channel Head/Neck Coil using the following imaging protocol: T1- weighted MP-RAGE images (FOV = 250 mm, 208 axial slices; 1 × 1 × 1 mm, TR = 1,200 ms, TE = 2.29 ms, TI = 600 ms), Flip angle = 8 degrees, 3D axial SWI images [FOV 220 mm, 88 slices/1.5 mm slice thickness (interleaved)/TR 27 ms, TE 29 ms, 1 average], DTI acquisition parameters were axial DTI/TA:10:14 min, FOV 256 mm, 70 slices, 2 mm slice thickness, TR 9,000 ms, TE 88 ms, Flip Angle 15°, 1 average, Acceleration Factor 2/ref. lines 24, diffusion directions 64, *b*-value 1:0 s/mm^2; b-value 2:1,000 s/mm^2. Diffusion images were also corrected for susceptibility distortions with the acquisition of a sequence with 64 PA reversed-phase directions. Double IR FLAIR images were obtained for the concussed patients (26).

For the other control patients ($n = 309$), Sagittal T1 FLAIR, Axial MP-RAGE, and DTI images were acquired identically to the concussed patients. All image sets in this study were anonymized and correlated with clinical data. Axial T2-FLAIR images were also obtained for control subjects.

Image analysis

Image preprocessing, segmentation, and quantification

For each subject, three MRI series (MPRAGE, Apparent Diffusion Coefficient, and Fractional Anisotropy) were co-registered using 3D Affine registration with spline deformation and segmented into two tissue types (gray and white matter) in eight anatomical subregions using a derivative of the 152c Montreal Brain atlas. These subregions included the right and left temporal, occipital, parietal, and frontal lobes. Five radiomic quantifications were performed, including raw signal measurement, fractal signature, and three-level wavelet decompositions, which provided information regarding the three-dimensional texture patterns of the measurements. The two tissues, three MRI series, eight subregions, and five radiomic measurements yielded 240 numeric sets, each further described by 33 statistical descriptions ranging from simple arithmetic mean and standard deviation to complex representations of numeric distribution, yielding a total set of 7,920 individual quantitative values for each subject. Logistic Regression with L1 regularization was used to estimate a subset of features to build the Post-Concussive Syndrome Index (PCSI) (34). The PCSI is a mathematical formula that produces a value between 0 and 1 for each subject based on the values of the subset of radiomic measurements, with values less than 0.5 regarded as consistent with uninjured healthy subjects and values of 0.5 or greater associated with post-concussive subjects. Thus, the PCSI value represents the probability of PCS based on the multidimensional MRI signal. Details of the image processing and ML are provided by Tamez-Peña et al. (26). All image processing was performed using CIPAS (Qmetrics Technologies, Rochester, NY), and prediction of the PCSI was carried out using R 4.1.2, with the FRESA.CAD 3.3.1 Package (35).

Statistical analysis

For each participant, we recorded the PCSI, sex, age, weight, health, origin of trauma or condition, and clinical symptoms at the time of the MRI. Age, weight, height, and PCSI of subjects in the validation set were described by mean and standard deviation and stratified by cases and controls. Statistical differences between the cases and controls were computed using t-tests for age, height, and weight. The PCSI was described using receiver operating characteristics (ROC) and the area under the ROC curve (ROCAUC). This paper also reports the accuracy, sensitivity, specificity, and diagnostic odds ratio between concussed and non-concussed subjects with their corresponding confidence intervals. The behavior of PCSI values for control group subsets (uninjured subjects with hearing loss, headaches, other indications, or young athletes) was described by frequency and PCSI association using violin plots and compared using ROC analysis. PCSI values of injured subjects were compared according to sex, loss of consciousness at the time of injury, mechanism of injury (sports, vehicular injuries, and assaults/falls), time interval from injury to MRI examination, prior concussion history, and reported symptoms. Significant associations with binary symptoms were computed using the Wilcoxon signed-rank test, and the effect size was described by ROCAUC. None of the *p*-values were adjusted for multiple comparisons. All statistical analyses were performed using R 4.2.2.

Results

Table 1 presents the demographic characteristics of the study participants. The MRI indication for the control/uninjured participants was hearing loss ($n = 132$), followed by headache ($n = 88$), and several other conditions ($n = 42$). The mechanism of injury in the case/PCS participants was sports-related ($n = 34$), motor vehicle use ($n = 14$), or other causes ($n = 21$). The sex distribution of the case/control sets was statistically different, where female participants were slightly more common in the control group than in the case group, although there were more females than males in both groups. There were marked differences between the age at injury of the participants (cases) and normal clinical patients (controls). Controls were slightly older (42.7 ± 11.2 years) than cases (31.9 ± 14.3 years). Another significant difference in this cohort was the weight of the control participants being heavier than the cases (90.3 ± 25.3 kg vs. 77.3 ± 20.7 kg, $p < 0.001$).

Regarding the PCSI, Figure 2 shows the distribution of the index across all patients stratified by sex. The PCSI could separate PCS subjects from non-injury subjects using the injury threshold

10.3389/fneur.2023.1282833

TABLE 1 Demographics of the PCSI validation cohort.

	All (n = 333)		Females (n = 189)		Males (n = 143)	
	Control (n = 264)	Case (n = 69)				
	(F/M/N) 150/113/1	(F/M/N) 39/30/0				
	MRI indication: Hear loss = 132 Headache = 88 Other = 42	Trauma origin: Sports = 34 MVA = 14 Other = 21	Control (n = 150)	Case (n = 39)	Control (n = 113)	Case (n = 30)
Age (years)	42.7 (11.2)	31.9 (14.3)***	41.6 (11.0)	33.1 (14.7)**	44.4 (11.2)	30.4 (13.8)***
Height (cm)	170.1 (10.4)	170.7 (10.9)	163.7 (7.3)	164.6 (7.7)	178.4 (7.6)	178.7 (9.3)
Weight (kg)	90.3 (25.3)	77.3 (20.7)***	82.7 (23.7)	71.6 (19.9)**	100.7 (23.7)	84.8 (19.6)**
PCS index	0.12 (0.17)	0.57 (0.34)***	0.13 (0.19)	0.57 (0.37)***	0.10 (0.14)	0.60 (0.31)***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

of 0.5 with an accuracy of 0.88 (95%CI [0.84, 0.92]), sensitivity of 0.64 (95%CI [0.51, 0.75]), specificity of 0.95 (95%CI [0.91, 0.97]), diagnostic odds ratio (DOR) of 31 (95%CI [15, 65]), and ROCAUC 0.87 (95%CI [0.82, 0.92]). The performance of the index was similar between males and females with sensitivity, specificity, and DOR of 0.63, 0.97, and 63 vs. 0.64, 0.93, and 23 for males and females, respectively.

Because there was a strong age imbalance between the cases and controls, we tested the hypothesis that the index was not associated with the participant's age. The test consisted of modeling the PCSI by participant age adjusted by subject class (Case or Control). The results of the model indicated that subject class predicted most of the variance in the PCSI ($p < 0.001$), while age did not ($p = 0.42$). Hence, the PCSI was not affected by participant age.

Exploratory analysis of the PCSI

Table 2 shows the distribution of the most common PCS symptoms and their association with the PCSI. The most common symptom of PCS was headache (87%), followed by concentration issues (62%). Headache was not associated with the index value, but the presence of vision problems was associated with a higher index value [0.671 vs. 0.478, $p(P > A) = 0.02$]. Concentration problems and anxiety showed a trend [$(P > A) < 0.1$] of higher index values when symptoms were present. Both the symptoms had a ROCAUC of 0.6. Figure 3 shows the distribution of the PCSI for the three top symptoms as well as the distribution of the PCSI and the presence of neck pain. The neck pain results showed a negative trend of having a lower index value for subjects reporting neck pain vs. patients without neck pain [$p = 0.440$, $A = 0.636$, $p(P < A) = 0.03$].

Figure 4 shows the distribution of the PCSI concerning the history of previous concussions, loss of consciousness, and the mechanism of trauma. In these cases, we observed a positive trend toward a higher PCSI value in the three exploratory analyses, but none of them reached statistical significance ($p > 0.1$).

We also explored the effect of elapsed time from injury on the PCSI and observed a non-significant trend of diminishing index values with increasing time from injury ($p = 0.12$).

The final exploratory analysis explored the behavior of the index in non-trauma participants. The results are presented at the bottom of Table 2. The hearing loss, headache, and other patient subcohorts had non-significantly different PCSI values than the other non-trauma patients [$p(A \neq B) = 0.16$, 0.88, and 0.09, respectively].

Discussion

This study showed that our previously developed ML-based classifier, the PCSI, enabled differentiation of real-world patients with clinically diagnosed mTBI from those without history of head injury with statistical accuracy (26). In this study, we further evaluated the performance of the PCSI by classifying a broader demographic of injured patients with PCS and comparing them with a broader control population of uninjured subjects aged between 18 and 60 years, imaged on multiple MRI scanners. The results of this study indicate that PCSI performed well in the real-world population, with a mean PCSI value of 0.57 for patients with PCS and a mean PCSI value of 0.12 for control subjects (ROCAUC 0.87, 95%CI [0.84, 0.92]). The PCSI had a sensitivity of 64%, specificity of 95%, and accuracy of 88% for the evaluation of patients with PCS. Clear differences in PCSI between controls and injured patients were observed across population subgroups, including females, males, sports-related mTBI, and non-sports-related mTBI, as well as sub-cohorts of injured and non-injured athletes. These results further suggest that the PCSI detects post-injury microtrauma, and that mTBI subjects with PCS exhibit significant differences in the content of the MRI information from combined structural and DTI image data as compared to uninjured subjects.

The PCSI performance of 88% accuracy with 65% sensitivity compares well with the reported work of Fleck et al. (36) that showed classification accuracy of only 62%. The main difference is that our work is based on an open adult population, while their study concentrated on an adolescent population. Other studies on the use of AI in mTBI showed similar performance, with sensitivity varying from 68% (Mitra) (28), 79% (Goswami) (29), to 89% (Vergara) (30). Even though these last three studies used mTBI subjects instead of PCS subjects, they illustrated that AI has an important role in identifying subtle differences in brain function or microstructure that may have clinical relevance in the near future.

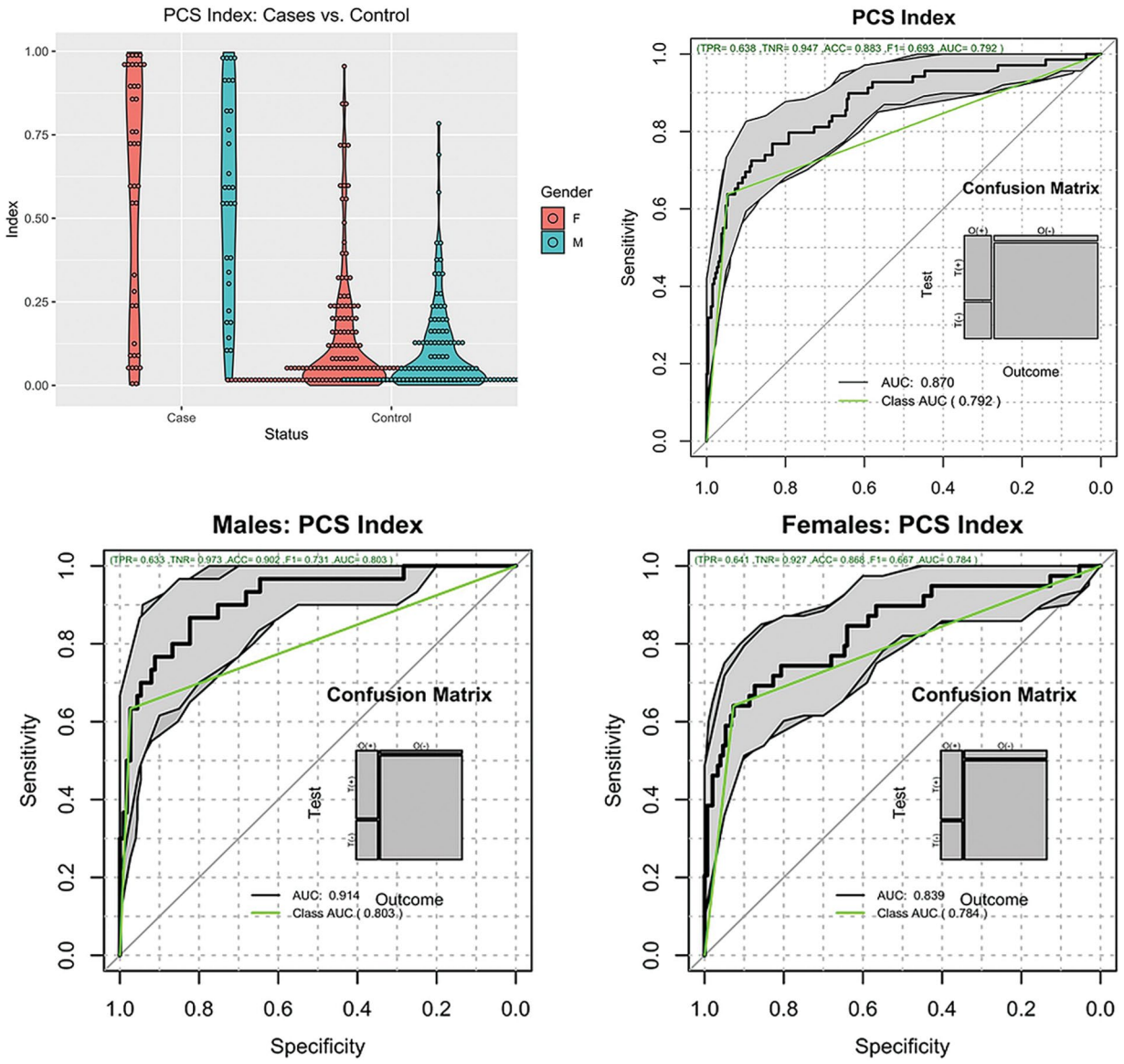


FIGURE 2
Distribution of the PCSI. Top left, violin plots of the cases and controls stratified by gender. Top right, ROC plot of all subjects. Bottom plots, ROC plots of Males and Females.

Although performance of the PSCI is good, the current method is based on a simple logistic model with L1 penalization. In the future, a more elaborate ML method may be able to be implemented, considering the source of individual misclassification errors, or the inclusion of additional clinically relevant information to further enhance accuracy. Furthermore, the presented results were validated using the same hardware, obtained using the same protocol on the same types of 3.0 Tesla MRI scanners, thus enabling data harmonization. Therefore, the performance of the developed PCSI must be evaluated on other types of MRI scanners, such as 3 or 1.5 Tesla, and with other diffusion tensor imaging protocols.

No significant differences were found in the PCSI based on sex for the control group or the mTBI cohort, consistent with previously reported results regarding the effect of sex on the frequency of PCS and rate of clinical recovery (37). The analysis of the PCSI showed a non-statistically significant, slightly downward trend associated with increased time intervals between injury and MRI. These findings are similar to those of multiple prior studies that showed persistent changes in individual DTI metrics from weeks to months after mTBI although the specific DTI metrics were inconsistent in their behavior (5, 18, 38–41). These observations are consistent with previous findings of partial resolution of alterations in AD and MD in cerebral white tracts from 2 weeks to 6 months post-mTBI, compared with persistent changes in lower FA. Contradictory results have been reported regarding the outcome of patients with PCS in relation to the number of prior concussions (17, 40, 42, 43). In our study, the relationship between the number of prior concussions and the PCSI showed a trend of increasing PCSI values for subjects with a history of more prior concussions; however, this did not reach statistical significance. In our study, no statistically significant differences were found in the PCSI values when comparing sub-cohorts of injured subjects who lost consciousness at the time of injury and those who did not, consistent with previous reports that loss of consciousness is not associated with an increased rate and duration of PCS (17, 44).

TABLE 2 Symptoms and PCSI.

	Symptom	Frequency (%)	Mean index present	Mean index absent	ROCAUC
Cases ($n = 69$)	Headache	60 (87%)	0.567	0.572	0.50
	Concentration	43 (62%)	0.604	0.507	0.59
	Photophobia	39 (57%)	0.535	0.610	0.43
	Fatigue	35 (51%)	0.593	0.542	0.56
	Sleep	35 (51%)	0.584	0.551	0.53
	Dizziness	34 (49%)	0.550	0.585	0.46
	Memory problems	34 (49%)	0.528	0.606	0.44
	Mood problems	33 (48%)	0.572	0.564	0.49
	Noise sensitivity	33 (48%)	0.532	0.600	0.44
	Vision problems	32 (46%)	0.671	0.478	**0.66***
	Anxiety	31 (45%)	0.633	0.514	0.60
	Nausea	31 (45%)	0.594	0.546	0.55
	Irritability	28 (41%)	0.543	0.585	0.54
	Neck pain	24 (35%)	0.440	0.636	0.34*
	Balance	21 (30%)	0.489	0.602	0.40
Controls ($n = 264$)	Hearing loss	132 (50%)	0.118	0.100	0.50
	Headache	88 (34%)	0.143	0.094	0.54
	Other	44 (16%)	0.147	0.121	0.44

Bold values indicate ROC AUC greater than 0.6. *$p < 0.1$.

The exploratory analysis of the case subjects revealed that certain symptoms were associated with higher PCSI values. Specifically, we found that participants with vision problems had higher PCSI values than those without vision problems. Similar trends were observed for anxiety and concentration. As illustrated in Figure 3, subjects who reported vision problems, including blurring, had higher index values than those who did not (45). Upon examination, all patients with vision problems complained of either binocular blurred vision or convergence insufficiency. At least 12 of the 30 participants with vision problems also reported LOC after a concussive episode. Unfortunately, a complete ophthalmologic report was unavailable for these subjects; therefore, our ability to localize the potential injury site was limited. However, blurred vision is often the result of diplopia or nystagmus, which can occur from dysfunction in pathways located in the midbrain, a region with known biomechanical vulnerability to concussive forces (45). The connection between these symptoms and midbrain dysfunction is more likely for the subset of subjects presenting with both blurred vision and LOC, indicating midbrain-thalami involvement.

It is worth noting that the PCSI was negatively correlated with the presence of neck pain in the subset of cases whose trauma originated from motor vehicle accidents, which suggests the possibility of PCS symptoms originating from neck injury rather than structural brain microtrauma, assuming greater prevalence of neck injury due to MVA than other causes of mTBI (46, 47).

Our prior study suggested that the PCSI was able to detect structural post-injury microtrauma in the brain, but was limited by a small, relatively young control demographic. In this study, we expanded the study of the PCSI to include subjects injured in everyday activities outside of athletics, and uninjured subjects undergoing neuroimaging for tinnitus and migraine, but otherwise healthy, with normal MRI exam findings and no history of head trauma. This larger study population, imaged on multiple MRI scanners, demonstrated that the PCSI continued to accurately differentiate subjects with history of mTBI and PCS from those without, even those with similar symptoms (headache). The results presented continue to suggest the PCSI detects structural microtrauma associated with mTBI. Significant limitations include a lack of objective data confirming neurological injury, such as histopathology or force vector data from the injury event that could be compared with PSCI results to show positive correlation with injury. PCSI values could add confidence to diagnosis of concussion and determination of return to activity based on resolution of symptoms. Although our data trends show the PCSI diminishes with time from injury, we have no proof that the PCSI follows patient recovery from mTBI, given the lack of PCSI information from mTBI patients without PCS, e.g., those who have clinically recovered. Thus, the results presented in this study do not conclusively indicate whether the PCSI can predict the resolution of PCS or whether brain changes observed due to mTBI are permanent fixtures in affected patients. We hope to address this in future work examining longitudinal correlation of PCSI values and symptomatic recovery, perhaps providing prognostic information about recovery timelines. Sub-regional PSCI information could also help physicians determine patient-specific recovery therapy.

Conclusion

The results of this study show that the previously developed PCSI applied to multiparametric MR data from individuals aged between

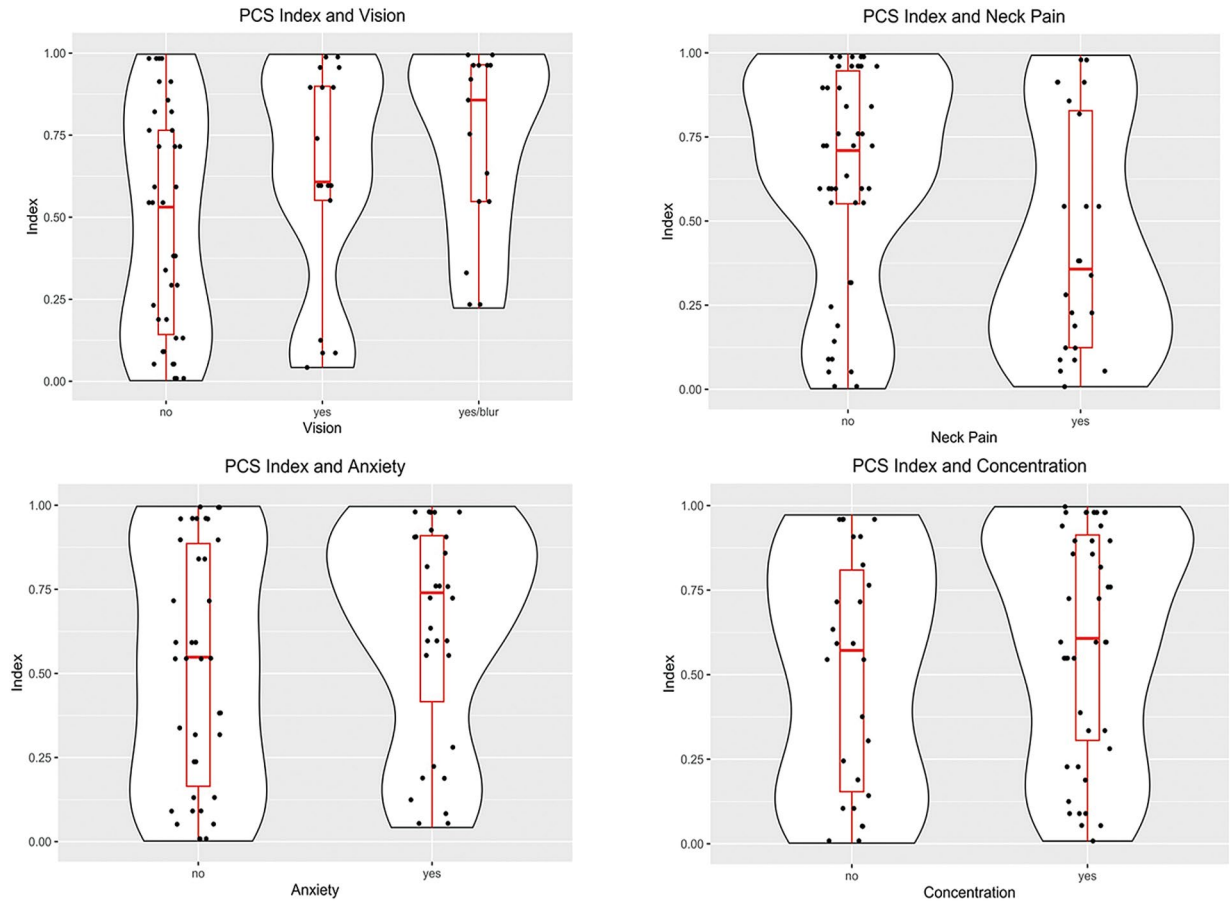


FIGURE 3
Violin plots of the distributions of the PCSI based on the case symptoms. Subjects reporting vision problems had PCSI values larger than subjects without problems ($p = 0.02$). Subjects reporting neck pain had lower PSCI values than subjects without neck pain ($p = 0.03$). The other symptoms had non-significant differences between the presence or absence of symptoms.

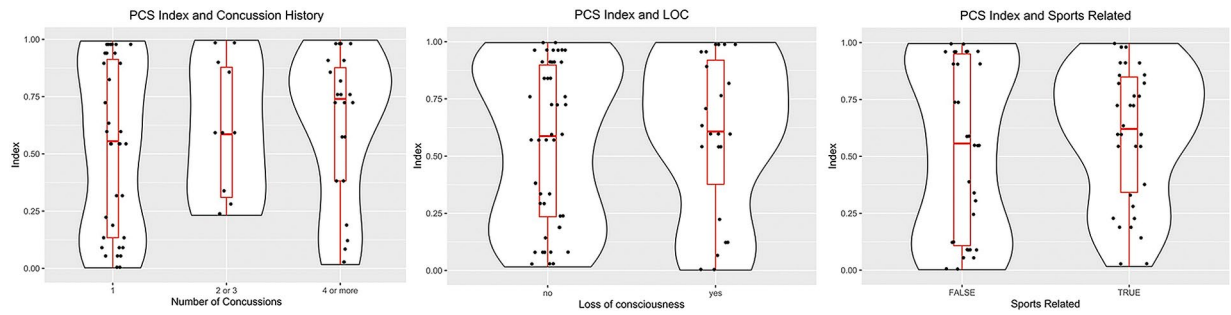


FIGURE 4
Violin plots showing the distribution of case subjects according to the history/origin of the concussion. Left, distribution based on the number of concussions. Middle, PCSI distribution according to the history of loss of consciousness. Right, differences between the origin of the trauma: Sports injury or another traumatic event.

18 and 60 years can accurately classify and differentiate patients with PCS from controls from 2 weeks to 1 year after mTBI with high sensitivity, specificity, and accuracy. No statistically significant differences were found in the PCSI values when compared by sex or loss of consciousness at the time of injury and those who did not. The results of this study suggest that the PCSI has great potential as an objective clinical tool to support the diagnosis, treatment, and follow-up care of patients with PCS. Further research is required to investigate the replicability of this method using other types of clinical MRI scanners. The PCSI could also provide sub-regional information about MRI-based structural abnormalities; additional investigation to compare localized PCSI data with objective data on injury localization would further increase confidence in the correlation of the PCSI values with post-traumatic brain injury microtrauma. Finally, work to

examine PCSI behavior on patients recovering from mTBI including those who have recovered symptomatically, e.g., without PCS, could provide useful information on the relationship between symptomatic resolution and whether this is due to brain plasticity or the healing of structural injuries.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the University of Rochester Office for Human Subject Protection. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because This research study involves a retrospective analysis of existing data only and there is no more than minimal risk to subjects. Due to the size of the study it would be impractical to locate and contact the potential subjects and burdensome for them to provide consent. In addition, no personally-identifiable information is being extracted, making it impossible to identify and locate individuals to request consent. Once subjects are selected for inclusion in this study, PHI will be removed for analysis and results reporting. For these reasons, waiving informed consent will not adversely affect the rights of subjects.

Author contributions

SM: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. AH: Writing – original draft, Writing – review & editing. PG: Formal analysis, Writing – review & editing. JB: Writing – review & editing. MM: Writing – review & editing. KR: Writing – review & editing. HM: Writing – review & editing. PR: Data curation, Writing – review & editing. ST: Formal analysis, Validation, Writing – original draft, Writing – review & editing. ES: Formal analysis, Project administration, Writing – original draft, Writing – review & editing. JT-P: Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to acknowledge Jared Noble for assistance with data handling and management, and Colleen Cottrell for assistance with coordinating IRB submission and manuscript review.

Conflict of interest

JT-P, ST, ES, and PG were shareholders of Qmetrics Technologies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Taylor CA, Bell JM, Breiding MJ, Xu LK. Traumatic brain injury-related emergency department visits, hospitalizations, and deaths - United States, 2007 and 2013. *MMWR Surveill Summ*. (2017) 66:1–16. doi: 10.15585/mmwr.ss6609a1

2. Langlois JA, Rutland-Brown W, Wald MM. The epidemiology and impact of traumatic brain injury - a brief overview. *J Head Trauma Rehabil*. (2006) 21:375–8. doi: 10.1097/00001199-200609000-00001

3. Narayana PA. White matter changes in patients with mild traumatic brain injury: MRI perspective. *Concussion*. (2017) 2:CNC35. doi: 10.2217/cnc-2016-0028

4. McMahon P, Hricik A, Yue JK, Puccio AM, Inoue T, Lingsma HF, et al. Symptomatology and functional outcome in mild traumatic brain injury: results from the prospective TRACK-TBI study. *J Neurotrauma*. (2014) 31:26–33. doi: 10.1089/neu.2013.2984

5. Asken BM, DeKosky ST, Clugston JR, Jaffee MS, Bauer RM. Diffusion tensor imaging (DTI) findings in adult civilian, military, and sport-related mild traumatic brain injury (mTBI): a systematic critical review. *Brain Imaging Behav*. (2018) 12:585–612. doi: 10.1007/s11682-017-9708-9

6. Levin HS, Diaz-Arrastia RR. Diagnosis, prognosis, and clinical management of mild traumatic brain injury. *Lancet Neurol*. (2015) 14:506–17. doi: 10.1016/S1474-4422(15)00002-2

7. Shenton ME, Hamoda HM, Schneiderman JS, Bouix S, Pasternak O, Rathi Y, et al. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. *Brain Imaging Behav*. (2012) 6:137–92. doi: 10.1007/s11682-012-9156-5

8. McCrory P, Feddermann-Demont N, Dvořák J, Cassidy JD, McIntosh A, Vos PE, et al. What is the definition of sports-related concussion: a systematic review. *Br J Sports Med*. (2017) 51:877–87. doi: 10.1136/bjsports-2016-097393

9. Vallières M, Kay-Rivest E, Perrin LJ, Liem X, Furstoss C, Aerts HJWL, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci Rep*. (2017) 7:10117. doi: 10.1038/s41598-017-10371-5

10. McCrory P, Meeuwisse WH, Aubry M, Cantu RC, Dvořák J, Echemendia RJ, et al. Consensus statement on concussion in sport: the 4th international conference on concussion in sport, Zurich, November 2012. *J Athl Train*. (2013) 48:554–75. doi: 10.1136/bjsports-2013-092313

11. Ruff RM, Iverson GL, Barth JT, Bush SS, Broshek DKthe NAN Policy and Planning Committee. Recommendations for diagnosing a mild traumatic brain injury: a National Academy of neuropsychology education paper. *Arch Clin Neuropsychol*. (2009) 24:3–10. doi: 10.1093/arclin/acp006

12. Carroll LJ, Cassidy JD, Holm L, Kraus J, Coronado VG. Methodological issues and research recommendations for mild traumatic brain injury: the WHO collaborating Centre task force on mild traumatic brain injury. *J Rehabil Med*. (2004) 36:113–25. doi: 10.1080/16501960410023877

13. Kontos AP, Elbin RJ, Schatz P, Covassin T, Henry L, Pardini J, et al. A revised factor structure for the post-concussion symptom scale baseline and postconcussion factors. *Am J Sports Med*. (2012) 40:2375–84. doi: 10.1177/0363546512455400

14. Scorza KA, Cote W. Current concepts in concussion: initial evaluation and management. *Am Fam Physician*. (2019) 99:426–34.

15. Mullally WJ. Concussion. *Am J Med*. (2017) 130:885–92. doi: 10.1016/j.amjmed.2017.04.016

16. McCrory P, Meeuwisse W, Dvorak J, Aubry M, Bailes J, Broglio S, et al. Consensus statement on concussion in sport-the 5th international conference on concussion in sport held in Berlin, October 2016. *Br J Sports Med*. (2017) 51:838–47. doi: 10.1136/bjsports-2017-097699

17. Tator CH, Davis HS, Dufort PA, Tartaglia MC, Davis KD, Ebraheem A, et al. Postconcussion syndrome: demographics and predictors in 221 patients. *J Neurosurg*. (2016) 125:1206–16. doi: 10.3171/2015.6.JNS15664

18. Gumus M, Santos A, Tartaglia MC. Diffusion and functional MRI findings and their relationship to behaviour in postconcussion syndrome: a scoping review. *J Neurol Neurosurg Psychiatry*. (2021) 92:1259–70. doi: 10.1136/jnnp-2021-326604

19. McCauley SR, Boake C, Levin HS, Contant CF, Song JX. Postconcussional disorder following mild to moderate traumatic brain injury: anxiety, depression, and social support as risk factors and comorbidities. *J Clin Exp Neuropsychol*. (2001) 23:792–808. doi: 10.1076/jcen.23.6.792.1016

20. Nelson LD, Temkin NR, Dikmen S. Recovery after mild traumatic brain injury in patients presenting to US level 1 trauma centers: a transforming research and clinical knowledge in traumatic brain injury (TRACK-TBI) study. *JAMA Neurol*. (2019) 76:1049–520. doi: 10.1001/jamaneurol.2019.1313

21. Ganesh A, Galetta S. Editors' note: cognitive outcome 1 year after mild traumatic brain injury: results from the TRACK-TBI study. *Neurology*. (2022) 99:773–3. doi: 10.1212/WNL.0000000000201414

22. Rausa VC, Shapiro J, Seal ML, Davis GA, Anderson V, Babl FE, et al. Neuroimaging in paediatric mild traumatic brain injury: a systematic review. *Neurosci Biobehav Rev*. (2020) 118:643–53. doi: 10.1016/j.neubiorev.2020.08.017

23. Levin HS, Temkin NR, Barber J, Nelson LD, Robertson C, Brennan J, et al. Association of sex and age with mild traumatic brain injury-related symptoms: a TRACK-TBI study. *JAMA Netw Open*. (2021) 4:e213046. doi: 10.1001/jamanetworkopen.2021.3046

24. Lui YW, Xue YY, Kenul D, Ge YL, Grossman RI, Wang Y. Classification algorithms using multiple MRI features in mild traumatic brain injury. *Neurology*. (2014) 83:1235–40. doi: 10.1212/WNL.0000000000000834

25. McDonald MA, Tayebi M, McGeown JP, Kwon EE, Holdsworth SJ, Danesh-Meyer HV. A window into eye movement dysfunction following mTBI: a scoping review of magnetic resonance imaging and eye tracking findings. *Brain Behav*. (2022) 12:e2714. doi: 10.1002/brb3.2714

26. Tamez-Peña J, Rosella P, Totterman S, Schreyer E, Gonzalez P, Venkataraman A, et al. Post-concussive mTBI in student athletes: MRI features and machine learning. *Front Neurol*. (2022) 12:12. doi: 10.3389/fneur.2021.734329

27. Luo XP, Lin DZ, Xia SW, Wang D, Weng X, Huang W, et al. Machine learning classification of mild traumatic brain injury using whole-brain functional activity: a radiomics analysis. *Dis Markers*. (2021) 2021:1–7. doi: 10.1155/2021/3015238

28. Mitra J, Shen KK, Ghose S, Bourgeat P, Fripp J, Salvado O, et al. Statistical machine learning to identify traumatic brain injury (TBI) from structural disconnections of white matter networks. *Neuroimage*. (2016) 129:247–59. doi: 10.1016/j.neuroimage.2016.01.056

29. Goswami R, Dufort P, Tartaglia MC, Green RE, Crawley A, Tator CH, et al. Frontotemporal correlates of impulsivity and machine learning in retired professional athletes with a history of multiple concussions. *Brain Struct Funct*. (2016) 221:1911–25. doi: 10.1007/s00429-015-1012-0

30. Vergara VM, Mayer AR, Damaraju E, Kiehl KA, Calhoun V. Detection of mild traumatic brain injury by machine learning classification using resting state functional network connectivity and fractional anisotropy. *J Neurotrauma*. (2017) 34:1045–53. doi: 10.1089/neu.2016.4526

31. Abdelrahman HAF, Ubukata S, Ueda K, Fujimoto G, Oishi N, Aso T, et al. Combining multiple indices of diffusion tensor imaging can better differentiate patients with traumatic brain injury from healthy subjects. *Neuropsychiatr Dis Treat*. (2022) 18:1801–14. doi: 10.2147/NDT.S354265

32. Nakagawa M, Nakaura T, Namimoto T, Kitajima M, Uetani H, Tateishi M, et al. Machine learning based on multi-parametric magnetic resonance imaging to differentiate glioblastoma multiforme from primary cerebral nervous system lymphoma. *Eur J Radiol*. (2018) 108:147–54. doi: 10.1016/j.ejrad.2018.09.017

33. Nakagawa M, Nakaura T, Yoshida N, Azuma M, Uetani H, Nagayama Y, et al. Performance of machine learning methods based on multi-sequence textural parameters using magnetic resonance imaging and clinical information to differentiate malignant and benign soft tissue tumors. *Acad Radiol*. (2023) 30:83–92. doi: 10.1016/j.acra.2022.04.007

34. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B Methodol*. (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x

35. Tamez-Pena JG, Tamez-Pena MJG. *Package 'FRESA. CAD'* (2014). Available at: https://rdocumentation.org/packages/FRESA.CAD/versions/3.4.7

36. Fleck DE, Ernest N, Asch R, Adler CM, Cohen K, Yuan W, et al. Predicting post-concussion symptom recovery in adolescents using a novel artificial intelligence. *J Neurotrauma*. (2021) 38:830–6. doi: 10.1089/neu.2020.7018. PMID: 33115345;

37. Gupte R, Brooks W, Vukas R, Pierce J, Harris J. Sex differences in traumatic brain injury: what we know and what we should know. *J Neurotrauma*. (2019) 36:3063–91. doi: 10.1089/neu.2018.6171

38. Kim E, Yoo RE, Seong MY, Oh BM. A systematic review and data synthesis of longitudinal changes in white matter integrity after mild traumatic brain injury assessed by diffusion tensor imaging in adults. *Eur J Radiol*. (2022) 147:110117. doi: 10.1016/j.ejrad.2021.110117

39. Jain B, Das AK, Agrawal M, Babal R, Purohit DK. Implications of DTI in mild traumatic brain injury for detecting neurological recovery and predicting long-term behavioural outcome in paediatric and young population-a systematic review. *Childs Nerv Syst*. (2021) 37:2475–86. doi: 10.1007/s00381-021-05240-6

40. Manley G, Gardner AJ, Schneider KJ, Guskiewicz KM, Bailes J, Cantu RC, et al. A systematic review of potential long-term effects of sport-related concussion. *Br J Sports Med*. (2017) 51:969–77. doi: 10.1136/bjsports-2017-097791

41. Wu YC, Harezlak J, Elsaid NMH, Lin Z, Wen Q, Mustafi SM, et al. Longitudinal white-matter abnormalities in sports-related concussion a diffusion MRI study. *Neurology*. (2020) 95:E781–92. doi: 10.1212/WNL.0000000000009930

42. Kerr ZY, Zuckerman SL, Wasserman EB, Vander Vegt CB, Yengo-Kahn A, Buckley TA, et al. Factors associated with post-concussion syndrome in high school student-athletes. *J Sci Med Sport*. (2018) 21:447–52. doi: 10.1016/j.jsams.2017.08.025

43. Gaudet CE, Iverson GL, Kissinger-Knox A, Van Patten R, Cook NE. Clinical outcome following concussion among college athletes with a history of prior concussion: a systematic review. *Sports Med Open*. (2022) 8:134. doi: 10.1186/s40798-022-00528-6

44. Zuckerman SL, Yengo-Kahn AM, Buckley TA, Solomon GS, Sills AK, Kerr ZY. Predictors of postconcussion syndrome in collegiate student-athletes. *Neurosurg Focus*. (2016) 40:E13. doi: 10.3171/2016.1.FOCUS15593

45. Hirad AA, Bazarian JJ, Merchant-Borna K, Garcea FE, Heilbronner S, Paul D, et al. A common neural signature of brain injury in concussion and subconcussion. *Sci Adv*. (2019) 5:eaau3460. doi: 10.1126/sciadv.aau3460

46. Kennedy E, Chapple C, Quinn D, Tumilty S. Can the neck contribute to persistent symptoms post concussion? Long-term follow up from a prospective descriptive case series. *J Man Manip Ther*. (2021) 29:318–31. doi: 10.1080/10669817.2021.1920276

47. Kasch H, Jensen LL. Minor head injury symptoms and recovery from whiplash injury: a 1-year prospective study. *Rehabil Process Outcome*. (2019) 8:117957271984563. doi: 10.1177/1179572719845634

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 19 December 2023
DOI 10.3389/fneur.2023.1294723

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Ziga Spiclin,
University of Ljubljana, Slovenia
Zhe Wang,
The First Affiliated Hospital of Guangdong Pharmaceutical University, China
Cheng Long,
South China Normal University, China

*CORRESPONDENCE
Hanwei Chen
✉ docterwei@sina.com;
✉ chenhanwei0756@pyhospital.com.cn
Yufeng Ye
✉ 838554325@qq.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 18 September 2023
ACCEPTED 30 November 2023
PUBLISHED 19 December 2023

CITATION
Liang J, Feng J, Lin Z, Wei J, Luo X, Wang QM, He B, Chen H and Ye Y (2023) Research on prognostic risk assessment model for acute ischemic stroke based on imaging and multidimensional data.
Front. Neurol. 14:1294723.
doi: 10.3389/fneur.2023.1294723

COPYRIGHT
© 2023 Liang, Feng, Lin, Wei, Luo, Wang, He, Chen and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Research on prognostic risk assessment model for acute ischemic stroke based on imaging and multidimensional data

Jiabin Liang[1,2,3†], Jie Feng[4†], Zhijie Lin[5†], Jinbo Wei[1], Xun Luo[6], Qing Mei Wang[7], Bingjie He[8], Hanwei Chen[1,3,8*] and Yufeng Ye[1,3*]

[1]Postgraduate Cultivation Base of Guangzhou University of Chinese Medicine, Panyu Central Hospital, Guangzhou, China, [2]Graduate School, Guangzhou University of Chinese Medicine, Guangzhou, China, [3]Medical Imaging Institute of Panyu, Guangzhou, China, [4]Radiology Department of Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [5]Laboratory for Intelligent Information Processing, Guangdong University of Technology, Guangzhou, China, [6]Kerry Rehabilitation Medicine Research Institute, Shenzhen, China, [7]Stroke Biological Recovery Laboratory, Spaulding Rehabilitation Hospital, Teaching Affiliate of Harvard Medical School, Charlestown, MA, United States, [8]Panyu Health Management Center, Guangzhou, China

Accurately assessing the prognostic outcomes of patients with acute ischemic stroke and adjusting treatment plans in a timely manner for those with poor prognosis is crucial for intervening in modifiable risk factors. However, there is still controversy regarding the correlation between imaging-based predictions of complications in acute ischemic stroke. To address this, we developed a cross-modal attention module for integrating multidimensional data, including clinical information, imaging features, treatment plans, prognosis, and complications, to achieve complementary advantages. The fused features preserve magnetic resonance imaging (MRI) characteristics while supplementing clinical relevant information, providing a more comprehensive and informative basis for clinical diagnosis and treatment. The proposed framework based on multidimensional data for activity of daily living (ADL) scoring in patients with acute ischemic stroke demonstrates higher accuracy compared to other state-of-the-art network models, and ablation experiments confirm the effectiveness of each module in the framework.

KEYWORDS

prognostic risk, assessment model, acute ischemic stroke, imaging, multidimensional data, MRI

1 Introduction

Stroke is a cerebrovascular disease characterized by localized cerebral ischemia, hypoxia leading to ischemic necrosis or softening, and subsequent neurological dysfunction. Approximately 16 million people worldwide suffer from stroke each year, with 5.7 million deaths and around 5 million disabilities (1). Survivors often experience difficulties in swallowing, speech impairment, motor dysfunction, cognitive impairment, emotional disorders, and other functional deficits (2, 3). Early diagnosis, prediction, and rehabilitation are key strategies for improving the prognosis of stroke patients. Stroke treatment guidelines emphasize that early diagnosis of stroke relies on imaging findings and clinical symptoms/signs. Neuroimaging plays a crucial role in the definitive diagnosis of suspected stroke patients.

However, limited studies have been conducted on the correlation between imaging-based predictions of complications in acute ischemic stroke, and most of them focus on single complications (4, 5). Several scales have been used clinically to predict functional outcomes in stroke patients (6, 7), such as the acute stroke registry and analysis of Lausanne (ASTRAL) and the totaled health risks in vascular events (THRIVE). However, these scales mostly incorporate variables at admission and are intended to provide information for treatment, without collecting post-treatment data for prediction. Therefore, in order to accurately assess the prognostic outcomes of patients, adjust treatment plans in a timely manner for those with poor prognosis, and intervene in modifiable risk factors, machine learning methods are needed to predict the prognostic risk of patients with acute ischemic stroke.

MRI is one of the crucial tools for evaluating acute ischemic stroke and has been widely used in clinical practice due to its high detection accuracy, sensitivity, and specificity. Computer-aided diagnosis (CAD) based on MRI has received extensive attention from researchers both domestically and internationally. For example, the texture analysis of apparent diffusion coefficient maps and diffusion-weighted imaging were used to predict the prognosis and subtype of ischemic stroke (8–10). A systematic review also demonstrated that a combined model combining clinical and imaging variables was more predictive of stroke outcome (11).

This study aimed to construct an acute ischemic stroke prognosis assessment model based on multidimensional imaging data, clinical information, treatment plans, prognosis, and complications, as shown in Figure 1. It contains the adaptive lesion awareness module (ALAM), the patient metadata encoder based on multilayer perceptron (MLP-FE), the cross-modal attention module (CMAM). This model provided a scientific basis for early clinical intervention, enabling healthcare professionals to make informed decisions and interventions based on the predicted prognosis.

2 Multidimensional ADL scoring network framework for acute ischemic stroke patients

In this section, we developed a multidimensional data-based cross-modal attention fusion network for the prognosis assessment of acute ischemic stroke patients. The structure of the network was illustrated in Figure 2 and consisted of three main components: the ALAM, the MLP-FE, and the CMAM, which were described as follows:

Firstly, the ALAM was based on a multi-scale global-local attention mechanism. It adaptively focused on the lesion region by learning joint features of global context information and local detailed features, enabling the extraction of more discriminative imaging features (shown as Figure 1).

Secondly, to fully utilize the patient metadata, we designed the MLP-FE. This encoder compiled the patient metadata into corresponding feature representations.

Lastly, the CMAM was proposed in this study for the fusion of multidimensional data, including clinical information, imaging features, treatment plans, prognosis, and complications. It aimed to achieve information complementarity and cross-validation by integrating data from different modalities.

2.1 Adaptive lesion sense module

In clinical practice, acute ischemic stroke patients exhibited complex MRI imaging features (shown in Figure 2): the location, size, and appearance of the lesions varied significantly among different patients and stages of development. Furthermore, the complications associated with stroke often had similar imaging characteristics, making them difficult to differentiate. Additionally, mild cases with smaller lesions were prone to being overlooked. This posed a significant challenge for deep learning models to learn lesion imaging

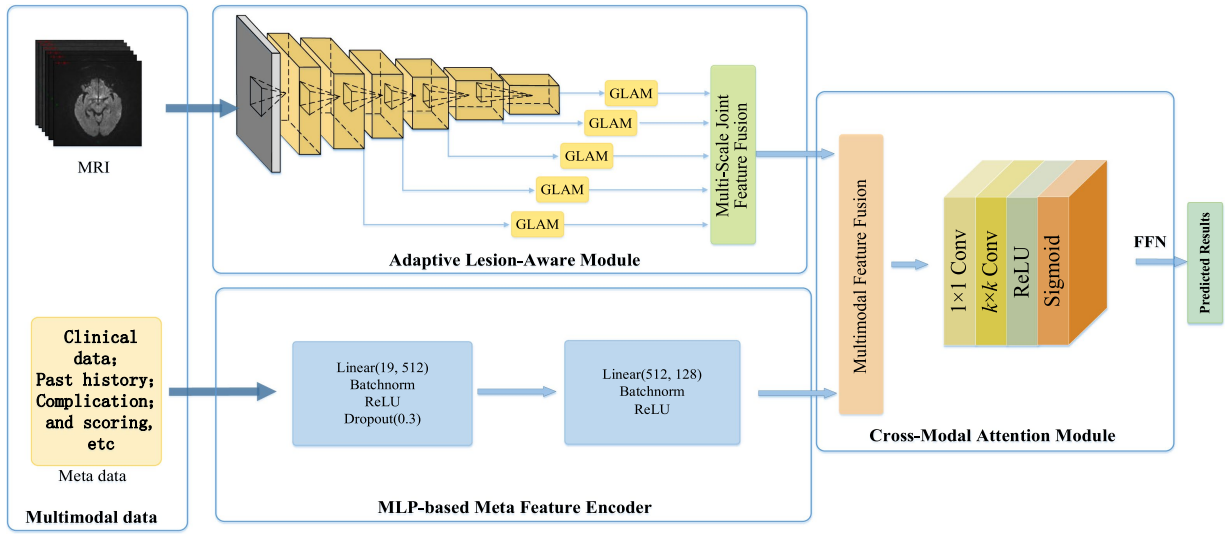


FIGURE 1
Based on the multi-dimensional network framework for the ADL score in patients with acute cerebral infarction.

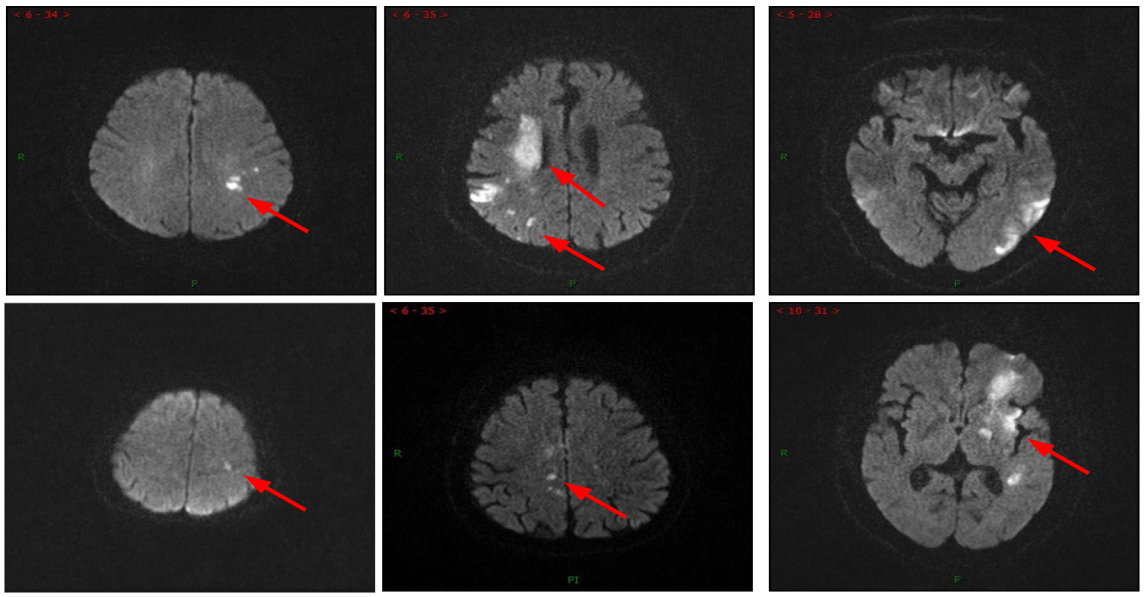


FIGURE 2
Example of the AIS patient MRI, where the red arrow indicates the focal area.

features from MRI, as it involved the issues of "same disease, different images" and "different diseases, same images."

In this subsection, we constructed an integration of the global-local attention module (GLAM) and the multiscale joint feature fusion module. This integration allowed for adaptive attention towards the relevant lesion regions and extraction of MRI images features by combining global and local views, akin to the process of radiologists' image interpretation. Additionally, the hierarchical design within this module facilitated multiscale joint feature fusion, enabling the extraction of lesion characteristics of different shapes and sizes. Moreover, these features interacted with each other to obtain the most effective image information.

2.1.1 The global-local attention module

The global-local attention module is used to extract joint features that capture both global context and local details. It consists of two attention blocks: the global attention block (GAB) based on self-attention mechanism and the local attention block (LAB) based on channel spatial attention. The GAB learns global context information to provide a comprehensive understanding of the scene, enabling the localization of lesion regions and the suppression of irrelevant background information. The LAB refines local features to capture more detailed lesion information, which helps in distinguishing similar imaging features of complications in acute ischemic stroke (AIS) and addressing the issue of "different diseases, same images."

Global context encompasses the implicit relationships between pixels and scene information in an image, providing a holistic perception of the scene. It has been widely used in various computer vision applications such as scene parsing and object detection (12–15). Intuitively, not all image content in MRI contributes to the final diagnosis, and irrelevant background information may even have a detrimental effect. However, by reasoning about the global scene, it becomes easier to detect and focus on the lesion region in the image. Existing methods primarily rely on convolutional neural networks (CNNs) for extracting MRI images features, which do not fully exploit and utilize the global context of the image. This is mainly due to the local nature of CNNs, which prevents them from learning the global context that aids in better lesion localization. Currently, effective modeling and integration of global context information in MRI images feature learning is an important research question that has not been fully investigated.

Inspired by natural language processing (NLP) networks, which extensively use transformers (14) to model global dependencies in language sequences, we designed a global attention block based on self-attention Transformer to model non-local interactions for learning global contextual information from local features. This global attention mechanism effectively detects lesion regions and efficiently disregards irrelevant background information. It is worth noting that, unlike segmentation-based methods, the GAB can locate lesion regions in an adaptive learning manner without any manual annotation. This enables more flexibility and robustness in handling various complex lesion shapes. The implementation process can be described as follows.

Taking into account that GAB required a sequence as input, similar to (15), the input feature map $F \in R^{C\times H\times W}$ was transformed into a sequence of flattened 2D patches $X_p \in R^{P^2C\times L}$, where (P, P) represented the resolution of each image patch, $L = HW/P^2$ was the number of image patches. Then, the image patches were mapped to a D-dimensional space through a learnable linear transformation, resulting in an output feature sequence $X \in R^{D\times L}$ which served as the input to GAB.

As shown in Figure 3, GAB consisted of two sub-layers: Multi-head self-attention (MHSA) and multi-layer perceptron (MLP). Each sub-layer was surrounded by residual connections, and layer normalization (LN) was applied before MHSA and MLP. Therefore, for an input feature sequence X, the output of GAB was given by

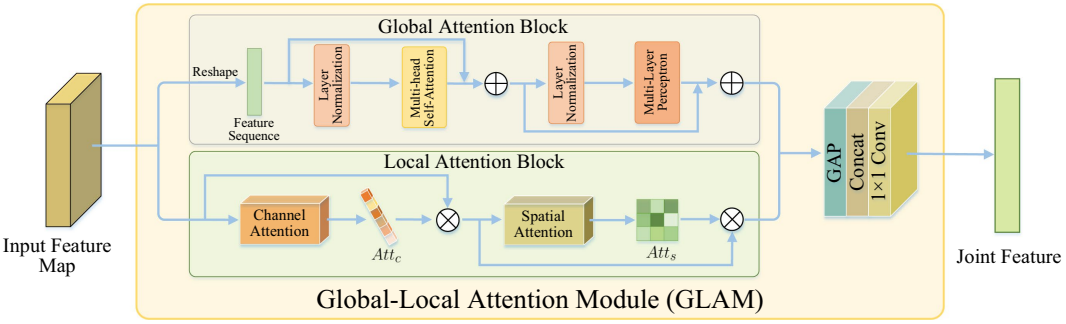


FIGURE 3
The structure of global-local attention module. Global attention block: learning global context information for detecting lesion regions; local attention block: refining local features to learn lesion details.

$$F_g = X' + \mathrm{MLP}\left[\mathrm{LN}\left(X'\right)\right], \tag{1}$$

among,

$$X' = X + \mathrm{MHSA}\left[\mathrm{LN}\left(X\right)\right]. \tag{2}$$

Although GAB can detect lesion regions through globally-guided attention, it is limited in learning lesion details. As the complications of AIS have similar imaging features, that is, there is a situation of "heterogeneous lesions and simultaneous shadows," which requires learning more detailed lesion features to distinguish them. However, traditional CNN is difficult to accurately extract subtle visual features from images (16). The features extracted by CNN contain two types of information: channels and spatial regions, which have different contributions to learning lesion details (17, 18). In order to improve the characteristics of network contribution to high sensitivity, make it concentrate more high-end features, in order to enhance expressing ability for the pathological change information. In this study, a LAB based on channel spatial attention was designed to conduct adaptive feature refinement and refinement by learning the relationship between channel and feature spatial location.

As shown in Figure 3, LAB was constructed sequentially from channels and spatial attention to effectively aid the flow of information within the network by learning which information to emphasize or suppress. Based on this, the LAB can be adaptive to refine local characteristic lesions in order to obtain more detailed information, and thus improve the network's ability to distinguish the similarity between classes of AIS complications.

Given a figure $F \in R^{C\times H\times W}$ as input, LAB sequentially output a channel attention map $\mathrm{A}tt_{\mathrm{c}} \in R^{C\times 1\times 1}$ and a 2D spatial attention map $\mathrm{A}tt_{\mathrm{s}} \in R^{1\times H\times W}$, which can be defined as

$$\begin{aligned} F' &= \mathrm{Att_c}\left(F\right) \circ F \\ F_{\mathrm{l}} &= \mathrm{Att_s}\left(F'\right) \circ F' \end{aligned} \tag{3}$$

In this context, $\circ$ represented the dot product, F' and F_{l} correspond to the channel branch and LAB output, respectively.

2.1.2 Multi-scale joint feature fusion module

Although many deep learning-based methods have achieved great success in stroke lesion detection, most of them rely on single-scale feature information and cannot fully utilize multi-scale feature representation, so it is difficult to deal with lesions of various scales and shapes, especially for small lesions (19, 20). This is because the following several pooling layer resolution is reduced, the characteristics of the small lesions may be lost. Therefore, this project designs a network with multi-scale feature information to realize the comprehensive extraction of MRI lesion features. It can combine the convolutional features of lesions of different scales and sizes together, retain low-level features of small lesions in the feature extraction process, and improve the diagnostic performance of early and mild patients.

In ALAM, a multi-scale feature extractor based on the pre-trained VGG16 (21) was first used to extract the feature representation of different layers. Then, at each scale, GLAM took the extracted feature map as input and learns joint features, implemented by two attention branches, capturing both global context and local refinement features. Finally, by integrating GLAM into the feature extractor, ALAM was designed to propagate joint features in a coarse-to-fine manner, fully extracted useful explicit feature representations at different scales, and exploit the complementary advantages of cross-scale implicit correlations. Therefore, the proposed network can effectively and robustly learn the imaging features of complex lesions by fusing multi-scale joint features, thereby significantly improving the detection performance of stroke lesions.

2.2 Based on the MLP patient metadata encoder

In clinical practice, radiologists will make medical diagnosis and evaluation of patients based on their imaging features and multi-dimensional clinical information database, including clinical data, past medical history, complications, etc. However, existing deep learning (DL) models only use a single imaging feature to predict the diagnosis of patients, which is not convincing and accurate. In view of this, in addition to imaging features, this project also established a multi-dimensional clinical information database, including clinical data, past medical history, complications, and self-care scores, which

can provide patients with higher quality, more accurate and more personalized medical diagnosis AIS.

In order to make full use of patient metadata, this study selected 19 metadata features from the multi-dimensional clinical information database of patients, and designed a patient metadata encoder based on MLP to compile the patient metadata into the corresponding feature representation. Nineteen metadata features included "hospital," "gender," "age ", "disorder ", "consciousness," "complications (hemiplegia) ", "pneumonia," "aphasia," "swallowing disorder," "facial paralysis," "dementia," "cognitive impairment," "depression," "first ability score," "hypertension," "diabetes," "atrial fibrillation," "coronary atherosclerotic heart disease," "hyperlipidemia," "high homocysteinemia." Specifically, there were two dense layers in this encoder, including 256 neurons. Each layer was followed by batch normalization and rectified linear unit (ReLU) activation layers. After the first dense layer attempts to generalize the metadata into the network, a further 25% was discarded using the dropout layer, thus avoiding possible overfitting of the data.

2.3 Cross-modal attention module

In order to strengthen the transfer between features and improve the performance of multimodal features, this study proposed a cross-modal attention module for multi-dimensional data information fusion such as clinical data, imaging features, treatment plans, prognosis and complications to achieve information complementarity and cross-validation. Multimodal fusion features not only retained the imaging features of MRI, but also made up for the clinical relevant information of patients, which provided more rich and comprehensive information for clinical diagnosis and treatment. The specific implementation process was as follows:

Firstly, the imaging features F_1 extracted by the adaptive lesion perception module and the metadata features F_2 obtained by the patient metadata encoder were combined into a multimodal joint feature $\boldsymbol{F}$:

$$F = \mathrm{Concat}\left(F_1, F_2\right). \tag{4}$$

Then, it was considered that the operation of directly concatenating these two features might propagate a large amount of useless information and noise generated during the encoding process to the decoding layer. Therefore, this project proposed to model the relationship between features by cross-modal attention to generate attention masks, and used it to adaptively select important features, which could inhibit the spread of some harmful information to a certain extent. The attention mask $\boldsymbol{A}$ can be obtained from the following formula:

$$\mathrm{A} = \sigma\left\{\mathrm{FC}\left\{\mathrm{ReLU}\left\{\mathrm{B}N\left[\mathrm{FC}\left(F\right)\right]\right\}\right\}\right\}, \tag{5}$$

in which, σ and **ReLU** were sigmoid and rectified linear unit activation functions, fully connected layer (**FC**) and batch normalization (**BN**) were fully connected layer and batch normalization, respectively.

Finally, the learned attention mask $\boldsymbol{A}$ was multiplied with the original feature map $\boldsymbol{F}$ to generate a feature map with attention weights F_o:

$$F_o = A \circ F, \tag{6}$$

in which, ∘ was the dot product. Based on this, the network can adaptively attention important characteristics, noise and suppress irrelevant information. These features were then fed into a classifier to predict the final outcome.

3 Experiment and result analysis

3.1 Data sets and experimental settings

3.1.1 Dataset

This study collected 337 patients diagnosed with acute cerebral infarction and included MRI examinations from January 2019 to January 2023 in Panyu Central Hospital of Guangzhou. Please refer to Table 1 and Supplementary material for details of the data set.

3.1.2 Experimental setup

This experiment was deployed in the PyTorch deep learning framework. The server used was equipped with two NVIDIA GeForce RTX 3090 GPUs with 24G memory. In the study, the input images are resize to 224×224 in both training and testing. During training, the cross-entropy loss function is applied to calculate the loss values between the predicted results and ground-truth labels. Moreover, the loss of the network is minimized by the Adam optimizer with a learning rate of 0.001, where the batch size is set to 32 and training is stopped after 200 epochs.

3.1.3 Evaluation indicators

This article selected the accuracy, precision, sensitivity (i.e., recall), $F1$-score as a model of evaluation index, the calculation formula, respectively, as follows:

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{7}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{8}$$

$$\mathrm{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{9}$$

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{10}$$

$$F1\text{–score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Sensitivity}}{\mathrm{Precision} + \mathrm{Sensitivity}} \tag{11}$$

TABLE 1 Distribution of cases (n = 337) and distribution of training test data (n = 3,106).

No.	Distribution pattern	Training set			Testing set		
		Eusemia	Poor prognosis	Total	Eusemia	Poor prognosis	Total
1.	Case distribution	142	95	237	60	40	100
2.	Training test data	1,136	977	2,113	497	396	893

Here, true positives (TP) and true negatives (TN) are the number of positive samples (that is, poor prognosis) and negative samples (that refers to good prognosis) that are correctly predicted, respectively. False positive (FP) is the number of negative samples misjudged as poor prognosis, and false negative (FN) is the number of poor prognosis samples misjudged as negative. In all experiments, the overall performance of the proposed and comparative methods was evaluated by calculating the mean and standard deviation of the cross-validation measures.

3.2 Experimental results

3.2.1 Comparison of performance indicators of different models

To validate the performance of the proposed model, it was compared with advanced image classification frameworks, including Res2Net (20), ResNet50 (22), CSPNet (23), EfficientNet (24), HRNet (25) and VGG16-GLAM. VGG16-GLAM is regarded as the proposed MRI analysis model that only takes MRI as input. The performance metrics of different models were shown in Table 1. The proposed model achieved an accuracy of 91.17%, precision of 87.26%, recall of 95.21%, and F1-score of 91.06%, which significantly outperformed existing CNN frameworks. Compared to ResNet50 (22), the proposed model showed improvements of 6.04% in accuracy, 20.97% in recall, and 4.65% in $F1$-score. Although our method did not achieve optimal precision, it significantly improved the prediction of adverse prognosis in stroke patients compared to other methods. This has important clinical implications in the treatment of stroke patients, and it confirmed that the fusion of multidimensional data information, including clinical data, radiological features, treatment plans, prognosis, and complications, was beneficial for assessing the quality of life in patients with acute ischemic stroke.

Furthermore, the Grad-CAM technique (26) was applied to visualize the proposed model. As observed in Figure 4, the Grad-CAM saliency maps of ResNet50 can roughly locate the lesion regions, but they may overlook or misjudge certain small lesions (e.g., the last two images in the last column). In contrast, the proposed model can accurately capture lesions of various complex shapes, even detecting and paying attention to minor abnormalities. This indicated that by learning joint features that encompass global context and local refinement, the model could adaptively detect relevant lesion regions to extract more discriminative feature representations, leading to better identification of patient case types, surpassing other state-of-the-art techniques. The experimental resulted in Figure 4 strongly support that the proposed model is intuitive and interpretable, confirming that the model's decisions primarily depend on the lesion regions while disregarding irrelevant image content.

3.2.2 Ablation experiments

In this subsection, we conducted a series of ablation experiments to validate the effectiveness of each key component of the prognostic assessment model for acute ischemic stroke based on multidimensional radiological data. Ablation studies for the proposed modules, including the multi-scale (MS) framework, GLAM, and the MLP. The No. 1 setting is the baseline network (i.e., VGG16) without these three modules, and then adding the to it as the No. 2 setting. GLAM is integrated into No. 2 to verify its effectiveness (No. 3). The No. 4 setting is the full version of the proposed network (see Table 2).

3.2.2.1 Effectiveness of the multiscale framework

To explore the contribution of the multiscale framework, we first used a single-scale network as the baseline model (No. 1) and compared it with other configurations. Table 3 showed that the proposed multiscale framework could learn better feature representations, resulting in performance improvements in all metrics compared to the single-scale framework. The No. 2 configuration using the multiscale framework achieved an accuracy improvement of 5.6%, sensitivity improvement of 12.12%, and precision improvement of 2.1%. The AUC of No. 1 and No. 2 was 0.843 and 0.894 respectively, and the model quality increased from 0.82 to 0.87. As seen in Figures 5A,B,E,F,I,J, the multiscale framework significantly enhanced the prediction of adverse prognosis in stroke patients. These results indicated that the multiscale framework was capable of capturing lesions of different scales and shapes, effectively leveraging the overall characteristics of brain MRI images.

3.2.2.2 Effectiveness of GLAM

Compared to No. 2, the inclusion of the GLAM module in No. 3 effectively improved the recognition performance, with an accuracy improvement of 1.35%, sensitivity improvement of 0.51%, and precision improvement of 2.37%. The AUC of No. 2 and No. 3 was 0.894 and 0.908 respectively, and the model quality increased from 0.88 to 0.89. As shown in Figures 5B,C,F,G,J,K, the addition of the GLAM module enhanced the negative prediction level for stroke patients compared to Figure 5J. This indicated that the GAB in GLAM models the global contextual information, constructing global-guided attention to adaptively focus on affected regions while disregarding irrelevant information for better lesion identification. The LAB could adaptively refine local features to obtain detailed lesion information. The combination of these two modules effectively utilized joint characteristics to adaptively focus on lesion regions and learned more detailed imaging features, thereby achieving better diagnostic performance. This could serve as the foundation for clinical prognosis assessment in stroke patients.

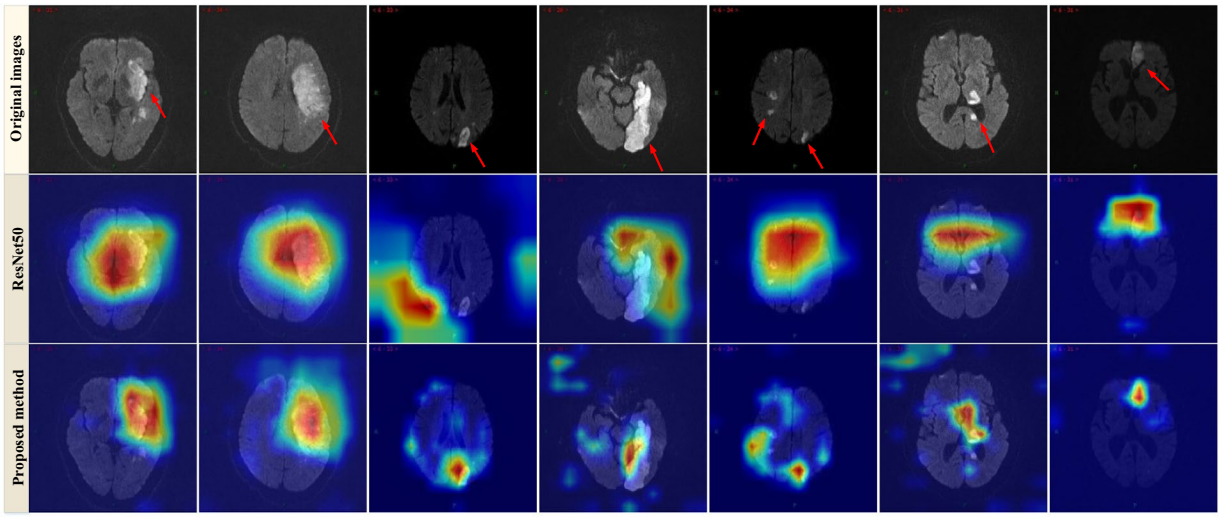


FIGURE 4
Examples of Grad-CAM thermal map of MRI. Original images (top row) and the Grad-CAM thermal map of the ResNet50 model (2nd row) and the proposed method (bottom row).

TABLE 2 Comparison of the performance indicators of the different models.

Method	Accuracy (%)	Precision (%)	Recall (%)	*F*1-score (%)
ResNet50	85.67 ± 1.12	91.87 ± 2.50	74.24 ± 1.76	82.12 ± 2.07
Res2Net	82.17 ± 2.77	82.24 ± 3.01	76.01 ± 3.28	79.00 ± 3.14
CSPNet	80.96 ± 2.35	82.10 ± 1.98	72.98 ± 2.77	77.27 ± 2.31
EfficientNet	82.08 ± 3.24	85.93 ± 2.75	70.96 ± 4.79	77.73 ± 3.49
HRNet	83.65 ± 2.79	84.72 ± 3.89	77.02 ± 2.52	80.68 ± 3.06
VGG16	83.87 ± 1.57	86.41 ± 2.89	75.50 ± 3.78	86.41 ± 3.28
VGG16-GLAM	90.82 ± 2.24	90.88 ± 2.86	88.13 ± 4.04	89.84 ± 3.34
Proposed method	91.71 ± 2.68	87.26 ± 1.39	95.21 ± 2.02	91.06 ± 1.58

3.2.2.3 Effectiveness of the patient metadata encoder based on MLP

Compared to No. 3, the fusion of patient metadata in No. 4 effectively improved the recognition performance, with an accuracy improvement of 0.89%, sensitivity improvement of 7.08%, and *F*1-score improvement of 1.22%. The AUC of No. 3 and No. 4 was 0.908and 0.916 respectively, and the model quality was 0.89. As observed in the comparison between Figures 5C,D,G,H,K,L, while there was a slight decrease in the negative prediction level, there was a significant improvement in the positive prediction level for stroke patients. In terms of clinical significance, the improved accuracy in positive prediction was more meaningful than negative prediction. This indicated that the inclusion of multidimensional clinical information could provide higher quality and more accurate assessments for patients.

4 Discussion

In the past few decades, various machine learning techniques, including logistic regression (LR) (27), linear discriminant analysis (LDA) (28), support vector machines (SVM) (29), decision trees (DT) (30), random forests (RF) (31), and neural networks (32), have been applied. These methods heavily rely on feature engineering, such as shape, texture, and pixel intensity distribution (histogram) obtained from computer programs, which can be used to identify potential imaging-based biomarkers and serve as input for improved machine learning models (33). SVM has improved the identification of carotid atherosclerosis (CA) from magnetic resonance brain images and prevented ischemic stroke in 97.5% of patients (34). The RF algorithm combined with geodesic active contour (GAC) model can automatically segment cerebrospinal fluid (CSF) in CT images for early identification of brain edema, a major medical complication after ischemic stroke (35). The LR algorithm has aided in the analysis of CT angiography (CTA) lesions and the discrimination of mobile intraluminal thrombus and atherosclerotic plaques, assisting in the selection of stroke treatment plans, with a sensitivity of 87.5% (36). The use of artificial neural networks to predict inadequate perfusion and the presence of effective collateral circulation in CT perfusion scans can facilitate further treatment, achieving an accuracy of 85.8% in testing on CT perfusion images of 396 patients (37). Several studies have employed machine learning methods on various public datasets to address various stroke-related issues for better improvement of healthcare systems and stroke treatment plans (38). However, the traditional machine learning approaches require preprocessing of input features and manual extraction. Optimization of image features and susceptibility to interference from multimodal imaging need further exploration and improvement (39). Recently, deep learning, as an emerging artificial intelligence (AI) technique, has the ability to automatically capture hierarchical and complex features from raw input data (40–42). Deep neural networks, with multiple layers, simulate the perception of the human brain, transforming "low-level"

TABLE 3 Effectiveness of component of the model.

No.	MS	GLAM	MLP	Accuracy	Precision	Recall	*F1*-score
1				83.87 ± 1.57	86.41 ± 2.89	75.50 ± 3.78	86.41 ± 3.28
2	√			89.47 ± 2.79	88.51 ± 1.02	87.62 ± 4.29	89.59 ± 1.65
3	√	√		90.82 ± 2.24	**90.88 ± 2.86**	88.13 ± 4.04	89.84 ± 3.34
4	√	√	√	**91.71 ± 2.68**	87.26 ± 1.39	**95.21 ± 2.02**	**91.06 ± 1.58**

Bold values indicate best performance.

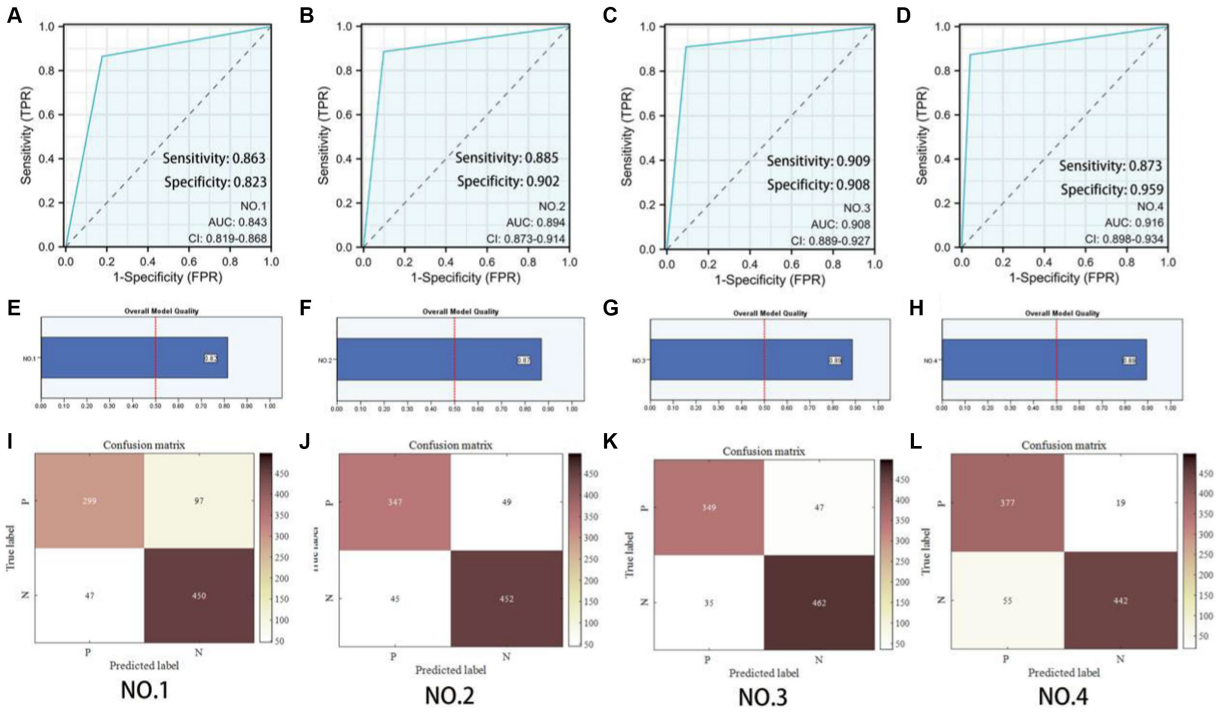


FIGURE 5
Predictive model evaluation. **(A–D)** Receiver Operating Characteristic curve. **(E–H)** Overall model quality. **(I–L)** Confusion matrixes of the ablation experiment. P: Poor prognosis; N: Eusemia.

to "high-level" representations, particularly in large-scale task solutions, in imaging classification, natural language processing, or bioinformatics (43, 44). In recent years, medical image processing has emerged as a hot research topic in deep learning, involving disease classification (45), lesion localization and segmentation, imaging reconstruction (46), and other tasks. Therefore, deep learning has been widely applied in stroke diagnosis and management, such as predicting clinical prognosis in AIS patients (47). Compared to traditional machine learning methods, deep CNN learning does not rely on handcrafted features. It automatically extracts and represents complex features when locating the core stroke lesions in CT or MRI (48). Deep learning not only saves time and effort but also captures pixel-level information of the lesions, contributing to improved diagnostic accuracy and prognosis (49).

Various artificial intelligence models have been widely applied in predicting the prognosis of ischemic stroke patients. Compared to predicting future stroke lesions on CT or MRI predicting patient prognosis is more challenging because commonly used prognostic scoring systems, such as the modified Rankin scale (mRS), are nonlinear and subjective, analyzing patients as a whole rather than on a voxel-by-voxel basis. This means there are fewer opportunities for artificial neural networks to learn from data, requiring larger training datasets to compensate for this limitation.

Previous studies mostly used non-imaging data as input and employed simple statistical models or machine learning models to predict prognosis (47, 50–53). However, CT or MRI can provide more information such as the size and location of infarctions. Tang et al. (54) utilized machine learning techniques combined with clinical data and the core-penumbra mismatch ratio from MRI and MRI perfusion to determine post-thrombolysis clinical outcomes. The short-term (7-day) result had an area under curve (AUC) of 0.863 [95% confidence interval (CI), 0.774–0.951], and the long-term (90 days) result had an AUC of 0.778 (95% CI, 0.668–0.888). Decision tree-based algorithms were able to predict the recovery outcomes (mRS >2 at 90 days) utilizing imaging and clinical data, with AUCs of 0.746 (extreme gradient boosting) and 0.748 (gradient boosting machine). Wang et al. (9) and Zhou et al. (8) used a multivariate logistic regression model to construct an imaging omics nomogram containing patient characteristics and imaging omics characteristics, and the AUC used to predict stroke outcome was greater than 0.80. Sun et al. (10) used clinical features and apparent diffusion coefficient maps to predict poor prognosis of acute stroke (mRS score >2) and

the AUC was 0.80. These models were superior to models using non-imaging data, and the clinical data were continuous and related, which demonstrates the great potential of the combination. The performance of this algorithm further improved when National Institute of Health stroke scale (NIHSS) 24 h and reperfusion status were included (55). Machine learning techniques, including regularized logistic regression, linear support vector machine, and random forest, outperformed existing pre-treatment scoring methods in predicting favorable clinical outcomes (90 days mRS >2) for patients undergoing thrombectomy for large vessel occlusion (LVO) (50).

Osama et al. (56) developed a parallel multi-parameter feature embedding Siamese network (PMFE-SN) that can learn from a small number of samples and handle skewness in multi-parameter MRI data. The proposed multi-parameter embedding architecture in PMFE-SN is based on deep learning and avoids overfitting even with a small number of samples in the dataset. The authors successfully predicted the prognosis of acute ischemic stroke patients 3 months later using MRI perfusion images and clinical data from the 2017 Ischemic Stroke LEsion Segmentation (ISLES) challenge, demonstrating superior performance compared to other advanced techniques.

Hilbert et al. (57) compared a deep learning model constructed using residual neural networks with a machine learning model utilizing traditional radiological image markers. The results showed that automatic image analysis using deep learning methods outperformed previous radiological image markers in predicting the prognosis of ischemic stroke patients and had the potential to improve treatment selection.

The proposed multi-dimensional ADL scoring network framework for AIS patients has higher accuracy than other state-of-the-art network models. Ablation experiments also confirmed the effectiveness of each module in the framework. In addition, the visualization results using the Grad-CAM technique show that our method can accurately locate the lesion area while ignoring irrelevant background information, indicating that the final identification results determined by the model are reliable and interpretable. This will help to provide more rich and comprehensive information for providing clinical diagnosis and treatment.

The effect of allowing machines to autonomously learn to fit nonlinear equations based on massive data rather than artificial formulas is closer to the complex problem itself. The same deep learning network can be trained for different types of samples and produce different fitting models individually to enhance its general applicability. Therefore, the development of acute cerebral infarction prognostic risk prediction models based on imaging and multi-dimensional data based on PyTorch deep learning framework is of great significance for early evaluation and intervention, guiding treatment plans and judging prognosis, reducing disability rate and reducing social and economic burden.

A total of 337 patients were included in our study, including 237 in the training set and 100 in the testing set. A study by Quan et al. (58), such as using fluid attenuated inversion recovery (FLAIR) and apparent diffusion coefficient (ADC) images to extract the image of omics characteristics to predict the prognosis of patients with AIS, included 190 cases of acute ischemic stroke patients, divided into the training group ($n = 110$) and external validation group ($n = 80$). In the study by Tang et al. (59) to predict the prognosis of patients with acute ischemic stroke receiving thrombolytic therapy, 168 patients with acute ischemic stroke were included. Compared with these studies on predicting the prognosis of ischemic stroke, our study not only included a large sample size, but also had a richer content, not limited to a specific treatment (thrombolysis, mechanical thrombectomy, etc.) and a single subtype of ischemic stroke. The larger sample size in our study provided more statistical power and enhances the reliability of the findings. This broader scope allowed for a more comprehensive understanding of the factors influencing the prognosis of ischemic stroke patients.

However, it is important to acknowledge the limitations of our study. Firstly, the retrospective nature of the study design introduces inherent limitations. Retrospective studies rely on existing data, which may be subject to selection bias and confounding factors. Prospective studies would provide more robust evidence and minimize potential biases. Secondly, all the samples used in our study were obtained from a single center, which might limit the generalizability of the findings. The patient population and treatment protocols at a single center may not be representative of other centers or populations. Therefore, multicenter validation is necessary to confirm the external validity and generalizability of our results.

To address these limitations, future studies could employ prospective designs with larger and more diverse samples, including patients from multiple centers. In addition, the inclusion of additional clinical and imaging variables could further enhance the predictive accuracy of our models. By addressing these limitations and conducting more rigorous studies, we can strengthen the evidence base and improve the prediction of prognosis in patients with ischemic stroke.

5 Conclusion

In this study, a model consisted of ALAM, MLP-FE, CMAM were constructed for multi-dimensional data information fusion, such as clinical data, imaging features, treatment plans, prognosis and complications, so as to achieve complementary advantages. The fusion features not only retain the MRI images features, but also make up for the clinical relevant information of the patient, which provided higher quality, more accurate and more personalized medical diagnostic for prognosis of AIS.

Data availability statememt

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Ethics statement

The Ethics Committee of Panyu District Central Hospital of Guangzhou reviewed and approved this study. The ethics committee waived the requirement of written informed consent for participation due to the retrospective nature of the study. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

JL: Data curation, Writing – original draft. JF: Methodology, Writing – original draft. ZL: Methodology, Writing – original draft. JW: Investigation, Writing – original draft. XL: Software, Writing – original draft. QW: Data curation, Writing – original draft. BH: Validation, Writing. HC: Project administration, Writing – original draft, Writing – review & editing. YY: Formal analysis, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Basic and Applied Basic Research Foundation of Guangdong Province (No. 2022A1515220217), Medical Scientific Research Foundation of Guangdong Province (Nos. A2023216 and A2022524), Science and Technology Program of Guangzhou (Nos. 202201010840, 202201010810, 202102080532, 202002030032, and 202002020023), Health Commission Program of Guangzhou (Nos. 20212A010025 and 20201A010085), Internal scientific research fund of Guangzhou Panyu Central Hospital (Nos. 2021Z001), Panyu General Science and Technology Plan Project (Nos. 2022-Z04-006), Panyu Major Science and Technology Plan Project (Nos. 2022-Z04-114).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2023.1294723/full#supplementary-material

References

1. Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. *Lancet Neurol*. (2007) 6:182–7. doi: 10.1016/S1474-4422(07)70031-5

2. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. (2018) 18:500–10. doi: 10.1038/s41568-018-0016-5

3. Panayides AS, Amini A, Filipovic ND, Sharma A, Tsaftaris SA, Young A, et al. AI in medical imaging informatics: current challenges and future directions. *IEEE J Biomed Health Inform*. (2020) 24:1837–57. doi: 10.1109/JBHI.2020.2991043

4. Ge Y, Wang Q, Wang L, Wu H, Peng C, Wang J, et al. Predicting post-stroke pneumonia using deep neural network approaches. *Int J Med Inform*. (2019) 132:103986. doi: 10.1016/j.ijmedinf.2019.103986

5. Li X, Wu M, Sun C, Zhao Z, Wang F, Zheng X, et al. Using machine learning to predict stroke-associated pneumonia in Chinese acute ischaemic stroke patients. *Eur J Neurol*. (2020) 27:1656–63. doi: 10.1111/ene.14295

6. Flint AC, Cullen SP, Rao VA, Faigeles BS, Pereira VM, Levy EI, et al. The THRIVE score strongly predicts outcomes in patients treated with the solitaire device in the SWIFT and STAR trials. *Int J Stroke*. (2014) 9:698–704. doi: 10.1111/ijs.12292

7. Papavasileiou V, Milionis H, Michel P, Makaritsis K, Vemmou A, Koroboki E, et al. ASTRAL score predicts 5 years dependence and mortality in acute ischemic stroke. *Stroke*. (2013) 44:1616–20. doi: 10.1161/STROKEAHA.113.001047

8. Zhou Y, Wu D, Yan S, Xie Y, Zhang S, Lv W, et al. Feasibility of a clinical-radiomics model to predict the outcomes of acute ischemic stroke. *Korean J Radiol*. (2022) 23:811–20. doi: 10.3348/kjr.2022.0160

9. Wang H, Sun Y, Ge Y, Wu PY, Lin J, Zhao J, et al. A clinical-radiomics nomogram for functional outcome predictions in ischemic stroke. *Neurol Ther*. (2021) 10:819–32. doi: 10.1007/s40120-021-00263-2

10. Sun Y, Zhuang Y, Zhu J, Song B, Wang H. Texture analysis of apparent diffusion coefficient maps in predicting the clinical functional outcomes of acute ischemic stroke. *Front Neurol*. (2023) 11:1132318. doi: 10.3389/fneur.2023.1132318

11. Dragoș HM, Stan A, Pintican R, Feier D, Lebovici A, Panaitescu PȘ, et al. MRI radiomics and predictive models in assessing ischemic stroke outcome—a systematic review. *Diagnostics*. (2023) 13:857. doi: 10.3390/diagnostics13050857

12. Lin Z, He Z, Xie S, Wang X, Tan J, Lu J, et al. AANet: adaptive attention network for COVID-19 detection from chest X-ray images. *IEEE Trans Neural Netw Learn Syst*. (2021) 32:4781–92. doi: 10.1109/TNNLS.2021.3114747

13. Shi M, Shen J, Yi Q, Weng J, Huang Z, Luo A, et al. LMFFNet: a well-balanced lightweight network for fast and accurate semantic segmentation. *IEEE Trans Neural Netw Learn Syst*. (2023) 34:3205–19. doi: 10.1109/TNNLS.2022.3176493

14. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al.. Attention is all you need, Proceedings of the 31st International Conference on Neural Information Processing Systems (2017) 6000–6010

15. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al.. An image is worth 16×16 words: transformers for image recognition at scale. International Conference on Learning Representations (2020)

16. Lin Z, He Z, Wang X, Zhang B, Liu C, Su W, et al. DBGANet: dual-branch geometric attention network for accurate 3D tooth segmentation. *IEEE Trans Circuits Syst Video Technol*. (2023):1. doi: 10.1109/TCSVT.2023.3331589

17. Mu N, Wang H, Zhang Y, Jiang J, Tang J. Progressive global perception and local polishing network for lung infection segmentation of COVID-19 CT images. *Pattern Recogn*. (2021) 120:108168. doi: 10.1016/j.patcog.2021.108168

18. Lin Z, He Z, Yao R, Wang X, Liu T, Deng Y, et al. Deep dual attention network for precise diagnosis of COVID-19 from chest ct images. *IEEE Trans Artif Intell*. (2022) 1:1–11. doi: 10.1109/TAI.2022.3225372

19. Yan Q, Wang B, Zhang W, Luo C, Xu W, Xu Z, et al. Attention-guided deep neural network with multi-scale feature fusion for liver vessel segmentation. *IEEE J Biomed Health Inform*. (2021) 25:2629–42. doi: 10.1109/JBHI.2020.3042069

20. Gao SH, Cheng MM, Zhao K, Zhang XY, Yang MH, Torr P. Res2Net: a new multi-scale backbone architecture. *IEEE Trans Pattern Anal Mach Intell*. (2021) 43:652–62. doi: 10.1109/TPAMI.2019.2938758

21. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition Proceedings of 3rd International Conference on Learning Representations (ICLR) (2015), pp. 1–14

22. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, (2016), pp. 770–778

23. Wang CY, Mark Liao HY, Wu YH, Chen PY, Hsieh JW, Yeh IH. CSPNet: a new backbone that can enhance learning capability of CNN. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2020. IEEE Computer Society. (2020)1571–1580

24. Tan M, Le Q. EfficientNet: rethinking model scaling for convolutional neural network. Proceedings of International Conference on Machine Learning (ICML). (2019): 6105–6114

25. Wang J, Sun K, Cheng T, Jiang B, Deng C, Zhao Y, et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans Pattern Anal Mach Intell*. (2021) 43:3349–64. doi: 10.1109/TPAMI.2020.2983686

26. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D Grad-CAM: visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE International Conference on Computer Vision. (2017): 618–626

27. Kleinbaum DG, Dietz K, Gail M, Klein M. *Logistic regression*. New York, NY: Springer-Verlag (2002).

28. Xanthopoulos P, Pardalos PM, Trafalis TB. *Robust data mining*. London: Springer (2013).

29. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. *IEEE Intell Syst Appl*. (1998) 13:18–28. doi: 10.1109/5254.708428

30. Rokach L, Maimon O. Decision trees In: *Data mining and knowledge discovery handbook*. Boston, MA: Springer (2005). 165–92.

31. Breiman L. Random forests. *Mach Learn*. (2001) 45:5–32. doi: 10.1023/A:1010933404324

32. Bishop CM. Neural networks and their applications. *Rev Sci Instrum*. (1994) 65:1803–32. doi: 10.1063/1.1144830

33. Castaneda-Vega S, Katiyar P, Russo F, et al. Machine learning identifies stroke features between species. *Theranostics*. (2021) 11:3017–34. doi: 10.7150/thno.51887

34. Bento M, Souza R, Salluzzi M, Rittner L, Zhang Y, Frayne R. Automatic identification of atherosclerosis subjects in a heterogeneous MR brain imaging data set. *Magn Reson Imaging*. (2019) 62:18–27. doi: 10.1016/j.mri.2019.06.007

35. Chen Y, Dhar R, Heitsch L, Ford A, Fernandez-Cadenas I, Carrera C, et al. Automated quantification of cerebral edema following hemispheric infarction: application of a machine-learning algorithm to evaluate CSF shifts on serial head CTs. *Neuroimage Clin*. (2016) 12:673–80. doi: 10.1016/j.nicl.2016.09.018

36. Thornhill RE, Lum C, Jaberi A, Stefanski P, Torres CH, Momoli F, et al. Can shape analysis differentiate free-floating internal carotid artery thrombus from atherosclerotic plaque in patients evaluated with CTA for stroke or transient ischemic attack? *Acad Radiol*. (2014) 21:345–54. doi: 10.1016/j.acra.2013.11.011

37. Vargas J, Spiotta A, Chatterjee AR. Initial experiences with artificial neural networks in the detection of computed tomography perfusion deficits. *World Neurosurg*. (2018) 124:e10–6. doi: 10.1016/j.wneu.2018.10.084

38. Sirsat MS, Fermé E, Câmara J. Machine learning for brain stroke: a review. *J Stroke Cerebrovasc Dis*. (2020) 29:105162. doi: 10.1016/j.jstrokecerebrovasdis.2020.105162

39. Sheth SA, Giancardo L, Colasurdo M, Srinivasan VM, Niktabe A, Kan P. Machine learning and acute stroke imaging. *J Neurointerv Surg*. (2023) 15:195–9. doi: 10.1136/neurintsurg-2021-018142

40. Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. *Adv Exp Med Biol*. (2020) 1213:3–21. doi: 10.1007/978-3-030-33128-3_1

41. Puttagunta M, Ravi S. Medical image analysis based on deep learning approach. *Multimed Tools Appl*. (2021) 80:24365–98. doi: 10.1007/s11042-021-10707-4

42. Alam M, Samad MD, Vidyaratne L, Glandon A, Iftekharuddin KM. Survey on deep neural networks in speech and vision systems. *Neurocomputing*. (2020) 417:302–21. doi: 10.1016/j.neucom.2020.07.053

43. Akkus Z, Cai J, Boonrod A, Zeinoddini A, Weston AD, Philbrick KA, et al. A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow. *J Am Coll Radiol*. (2019) 16:1318–28. doi: 10.1016/j.jacr.2019.06.004

44. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. (2015) 521:436–44. doi: 10.1038/nature14539

45. Zhang J, Xie Y, Wu Q, Xia Y. Medical image classification using synergic deep learning. *Med Image Anal*. (2019) 54:10–9. doi: 10.1016/j.media.2019.02.010

46. Kim M, Yun J, Cho Y, et al. Deep learning in medical imaging. *Neurospine*. (2020) 17:471–2. doi: 10.14245/ns.1938396.198.c1

47. van Os HJA, Ramos LA, Hilbert A, van Leeuwen M, van Walderveen MAA, Kruyt ND, et al. Predicting outcome of endovascular treatment for acute ischemic stroke: potential value of machine learning algorithms. *Front Neurol*. (2018) 9:784. doi: 10.3389/fneur.2018.00784

48. Li L, Wei M, Liu B, Atchaneeyasakul K, Zhou F, Pan Z, et al. Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. *IEEE J Biomed Health Inform*. (2021) 25:1646–59. doi: 10.1109/JBHI.2020.3028243

49. Zhang S, Zhang M, Ma S, Wang Q, Qu Y, Sun Z, et al. Research progress of deep learning in the diagnosis and prevention of stroke. *Biomed Res Int*. (2021) 2021:5213550–5. doi: 10.1155/2021/5213550

50. Nishi H, Oishi N, Ishii A, Ono I, Ogura T, Sunohara T, et al. Predicting clinical outcomes of large vessel occlusion before mechanical thrombectomy using machine learning. *Stroke*. (2019) 50:2379–88. doi: 10.1161/STROKEAHA.119.025411

51. 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) 50:1263–5. doi: 10.1161/STROKEAHA.118.024293

52. 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) 78:1916–22. doi: 10.1212/WNL.0b013e318259e221

53. Ho KC, Speier W, El-Saden S, Liebeskind DS, Saver JL, Bui AA, et al. Predicting discharge mortality after acute ischemic stroke using balanced data. *AMIA Annu Symp Proc*. (2014) 2014:1787–96.

54. Tang TY, Jiao Y, Cui Y, Zeng CH, Zhao DL, Zhang Y, et al. Development and validation of a penumbra-based predictive model for thrombolysis outcome in acute ischemic stroke patients. *EBioMedicine*. (2018) 35:251–9. doi: 10.1016/j.ebiom.2018.07.028

55. Xie Y, Jiang B, Gong E, Li Y, Zhu G, Michel P, et al. JOURNAL CLUB: use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information. *AJR Am J Roentgenol*. (2019) 212:44–51. doi: 10.2214/AJR.18.20260

56. Osama S, Zafar K, Sadiq MU. Predicting clinical outcome in acute ischemic stroke using parallel multi-parametric feature embedded Siamese network. *Diagnostics*. (2020) 10:858. doi: 10.3390/diagnostics10110858

57. Hilbert A, Ramos LA, van Os HJA, Olabarriaga SD, Tolhuisen ML, Wermer MJH, et al. Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke. *Comput Biol Med*. (2019) 115:103516. doi: 10.1016/j.compbiomed.2019.103516

58. Quan G, Ban R, Ren JL, Liu Y, Wang W, Dai S, et al. FLAIR and ADC image-based radiomics features as predictive biomarkers of unfavorable outcome in patients with acute ischemic stroke. *Front Neurosci*. (2021) 15:730879. doi: 10.3389/fnins.2021.730879

59. Tang TY, Jiao Y, Cui Y, Zhao DL, Zhang Y, Wang Z, et al. Penumbra-based radiomics signature as prognostic biomarkers for thrombolysis of acute ischemic stroke patients: a multicenter cohort study. *J Neurol*. (2020) 267:1454–63. doi: 10.1007/s00415-020-09713-7

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 13 March 2024
DOI 10.3389/fneur.2024.1343423

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Wenting Guo,
Capital Medical University, China
Qian Chen,
Capital Medical University, China

*CORRESPONDENCE
Jing Zhang
✉ ery_zhangjing@lzu.edu.cn

†These authors have contributed equally to this work

RECEIVED 23 November 2023
ACCEPTED 21 February 2024
PUBLISHED 13 March 2024

CITATION
Han N, Hu W, Ma Y, Zheng Y, Yue S, Ma L, Li J and Zhang J (2024) A clinical-radiomics combined model based on carotid atherosclerotic plaque for prediction of ischemic stroke.
Front. Neurol. 15:1343423.
doi: 10.3389/fneur.2024.1343423

COPYRIGHT
© 2024 Han, Hu, Ma, Zheng, Yue, Ma, Li and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A clinical-radiomics combined model based on carotid atherosclerotic plaque for prediction of ischemic stroke

Na Han[1,2,3†], Wanjun Hu[1,2,3†], Yurong Ma[1,2], Yu Zheng[1,2,3], Songhong Yue[1,2,3], Laiyang Ma[1,2,3], Jie Li[1,2] and Jing Zhang[1,2*]

[1]Department of Magnetic Resonance, Lanzhou University Second Hospital, Lanzhou, China, [2]Gansu Province Clinical Research Center for Functional and Molecular Imaging, Lanzhou, China, [3]Second Clinical School, Lanzhou University, Lanzhou, China

Objectives: To accurately predict the risk of ischemic stroke, we established a radiomics model of carotid atherosclerotic plaque-based high-resolution vessel wall magnetic resonance imaging (HR-VWMRI) and combined it with clinical indicators.

Materials and methods: In total, 127 patients were finally enrolled and randomly divided into training and test cohorts. HR-VWMRI three-dimensional T1-weighted imaging (T1WI) and contrast-enhanced T1WI (T1CE) were collected. A traditional model was built by recording and calculating radiographic features of the carotid plaques and patients' clinical indicators. After extracting radiomics features from T1WI and T1CE images, the least absolute shrinkage and selection operator (LASSO) algorithm was used to select the optimal features and construct the radiomics_T1WI model and the radiomics_T1CE model. The traditional and radiomics features were used to build combined models. The performance of all the models predicting ischemic stroke was evaluated in the training and test cohorts, respectively.

Results: Body mass index (BMI) and intraplaque hemorrhage (IPH) were independently related to ischemic stroke and were used to build the traditional model, which achieved an area under the curve (AUC) of 0.79 versus 0.78 in the training and test cohorts, respectively. The AUC value of the radiomics_T1WI model is the lowest in the training and test cohorts, but the prediction performance is significantly improved when the model combines IPH and BMI. The AUC value of the combined_T1WI model was 0.78 and 0.81 in the training and test cohorts, respectively. In addition, in the training and test cohorts, the radiomics_T1CE model based on HR-VWMRI combined clinical characteristics, which is the combined_T1CE model, had the highest AUC value of 0.84 and 0.82, respectively.

Conclusion: Compared with other models, the radiomics_T1CE model based on HR-VWMRI combined clinical characteristics, which is a combined_T1CE model, can accurately predict the risk of ischemic stroke.

KEYWORDS

stroke, carotid, atherosclerotic plaques, radiomics, magnetic resonance imaging, high-resolution vessel wall

Introduction

In 2015, the global mortality rate of ischemic cerebrovascular disease rose to second place among all causes of death. At the beginning of 2017, the China Stroke Epidemiology Survey team reported on the Circulation that the prevalence rate of stroke in China reached 1148.3/100000 after age standardization, of which ischemic stroke accounted for more than 70%, which has become a major disease seriously threatening the health of Chinese people (1–3). Atherosclerosis, a major cause of ischemic stroke, is a chronic progressive disease characterized by the atherosclerotic plaque formation. Embolism caused by carotid atherosclerotic plaque shedding has been recognized as accounting for approximately 18 to 25% of all strokes (4). In 2018, the American Society of Neuroradiology (ASNR) Vessel Wall Imaging Study Group published guidelines highlighting that the risk and severity of stroke associated with carotid plaques are not only related to the extent of luminal stenosis, but also related to plaque characteristics (5). An increasing amount of evidence indicates that vulnerable plaques are highly likely to lead to ischemic stroke and thrombotic complications, independent of the extent of luminal stenosis (6, 7). High-resolution vessel wall magnetic resonance imaging (HR-VWMRI) can not only assess luminal stenosis, but also characterize plaque morphology and different atherosclerotic components and identify vulnerable plaques (8). Plaque vulnerability imaging features include intraplaque hemorrhage (IPH), lipid-rich necrotic cores (LRNC) and thin fibrous caps, plaque inflammation, intraplaque neovascularization, plaque surface ulceration, and positive vascular remodeling (9). Randomized clinical trials have shown that HR-VWMRI is the most suitable and cost-effective imaging technique for characterizing plaque vulnerability characteristics. However, there are still drawbacks. First, HR-VWMRI has a longer scanning time, and image quality is sensitive to motion. Second, identifying vulnerable plaques based on HR-VWMRI is qualitative and subjective, and the results are greatly influenced by the personal factors of the researchers. Third, there is currently a lack of multicenter, large-scale, high-quality research on the relationship between vulnerable plaque characteristics and the risk of ischemic stroke. However, the relationship between carotid atherosclerotic plaque and ischemic stroke should go beyond the assessment of the primary imaging characteristics of plaque or the degree of vascular stenosis and adopt a new model in which carotid plaque imaging combined with artificial intelligence.

As an emerging multidisciplinary research field, radiomics integrates digital imaging information, statistics, artificial intelligence, machine learning, and deep learning methods to convert medical images into high-throughput quantitative features for research. Radiomics has been applied to disease diagnosis, tumor staging or grading, gene prediction, therapeutic effect evaluation, and prognosis judgment and plays an important role in assisting clinical decision-making. At present, radiomics studies on atherosclerotic plaques focus on identifying vulnerable plaques (10–12). The preliminary radiomics study of atherosclerotic plaque is based on texture analysis of CT or US images (13, 14). Compared with CT and US, VW-HRMRI has advantages in high soft tissue contrast and multisequence imaging, which can provide more valuable information. Shi et al. (11) conducted a radiomic study that was based on VW-HRMRI to distinguish stable and vulnerable basilar artery plaques. Subsequently, Shi et al. (15) performed a histogram texture analysis that was based on VW-HRMRI to extract the first-order texture features of atherosclerotic plaques in the middle cerebral artery and basilar artery, and they explored the differences in histogram features between stable and vulnerable plaques. Zhang et al. (16) established a high-risk carotid plaque model based on MRI radiomic features and evaluated its performance in distinguishing stable and vulnerable carotid plaques relative to the model based on traditional MRI features. It can be seen that the radiomic model showed better performance than the traditional model. These results indicate that compared with the traditional subjective qualitative and quantitative imaging characteristics, the radiomic method enables finding more differential features, showing its higher value in determining plaque vulnerability. This study established a radiomics model of carotid plaque based on HR-VWMRI and combined it with clinical indicators to accurately predict the risk of ischemic stroke.

Materials and methods

Study population

A total of 182 patients with a carotid plaque were consecutively recruited in this study between January 2020 and March 2022 in our hospital. All of them underwent HR-VW MRI 3D T1WI, T1CE, and head MRI (T1WI, T2WI, DWI, FLAIR, and 3D TOF-MRA). The study was approved by the ethics review board of our hospital (scientific research project ethics approval number: 2022A-695).

The exclusion criteria were as follows: (a) patients with cardiogenic stroke; (b) intracranial atherosclerosis, branch atheromatous disease, other causes (such as dissection, vasculitis, and vascular malformations), and causes remain unknown on ipsilateral infarction; (c) primary intracranial diseases; and (d) poor image quality. In total, 55 patients were excluded because of intracranial atherosclerosis, branch atheromatous disease, cardiogenic stroke and dissection on ipsilateral infarction ($n = 31$), vasculitis ($n = 16$), and poor image quality ($n = 8$), and 127 patients were finally enrolled in our study (flowchart in Figure 1). All patients were divided into stroke and stroke-free groups according to whether acute/subacute stroke (carotid territory) was shown on head MRI. We recorded the demographic and clinical characteristics of all participants including gender, age, body mass index (BMI), hypertension, hyperglycemia, hyperlipidemia, hyperuricemia, and hyperhomocysteinemia. Eligible patients were randomly divided into a training cohort ($n = 75$) and a test cohort ($n = 52$) with a ratio of approximately 3: 2.

MRI acquisition

The HR-VWMRI and head MRI were performed on a 3.0T MR scanner with a 20-channel head-and-neck phased array coil. Carotid MRI sequences include 3D TOF-MRA, 3D T1WI, T2WI, and 3D T1CE. Head MRI sequences include T1WI, T2WI, DWI, FLAIR, and 3D TOF-MRA. Detailed scanning parameters are listed in Table 1. The total scanning time was approximately 29 min.

Image analysis and segmentation

Independently identified and measured radiological features of the carotid plaques including IPH, LRNC, disrupted surface,

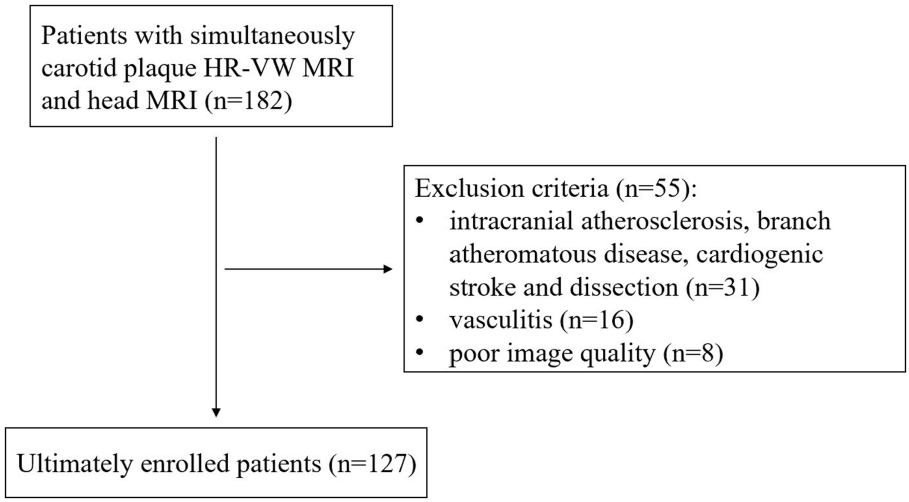


FIGURE 1
Study inclusion and exclusion flowchart.

TABLE 1 Detailed scanning parameters for the MRI protocol.

Sequence	Carotid plaque MRI				Head MRI				
	TOF	T1WI	T2WI	T1CE	T1WI	T2WI	DWI	FLAIR	TOF
FOV (mm)	220×128	240×240	160×160	240×240	230×230	230×230	230×230	230×230	230×230
Matrix	292×961	344×344	268×255	344×344	296×167	288×288	152×122	256/173	400×242
Slice thickness (mm)	1.2	0.8	3	0.8	5.5	5.5	5.5	5.5	1.2
Resolution (mm^2)	0.6×1.0	0.7×0.7	0.6×0.6	0.7×0.7	1.0×1.2	0.6×0.6	1.5×1.9	0.9×0.9	0.6×1.0
TR/TE (ms)	15/3.5	500/26	3500/60	500/26	2000/20	4000/97	2887/97	7500/110	22/3.5
NEX	1	1	2	1	1	1	1	1	1
Bandwidth (Hz/pixel)	216	637	273	637	159	439	1448	514	109
Flip angle (°)	18	75	90	75	90	90	90	90	18
Acquisition time	3min10s	6min46s	4min05s	6min46s	1min30s	2min24s	30s	1min21s	2min47s

TOF, time of flight; DWI, diffusion-weighted imaging; FLAIR, fluid-attenuated inversion recovery; FOV, field of view; TR, repetition time; TE, echo time; NEX, number of excitations.

enhancement, remodeling pattern, maximum wall area (Max WA), normalized wall index (NWI), and degree of stenosis by two senior radiologists (N.H. and YR.M.) with 8 or more years of experience in plaque imaging. If there are differences, the two radiologists reached a consensus after an additional reading session. Qualitative analysis of carotid plaques was performed on Picture Archiving and Communication Systems (PACS). For quantitative analysis of carotid plaques, VesselMass software (MEDIS, Version: 2014-EXP), a semi-automatic image analysis tool, was used. The slice with the largest plaque area was chosen, and the outer wall boundaries and inner lumen were manually outlined for measuring and calculating the total vessel area (TVA), the minimum luminal area (Min LA), the maximum wall area (Max WA), the normalized wall index (NWI), the remodeling index (RI), and the degree of stenosis. The relevant calculation formulas were as follows: (1) Max WA = TVA-Min LA (17); (2) NWI = Max WA/TVA (18, 19); (3) RI = vessel area at the point of maximum stenosis/reference vessel area at the distal portion (7); (4) Stenosis = (normal diameter at the distal portion-narrowest diameter at the stenosis)/normal diameter at the distal portion (20). If the RI is greater than 1, it is defined as positive remodeling; otherwise, it is defined as negative remodeling.

The open-source software ITK-SNAP (version 3.8.0)[1] was used to plaque segmentation for radiomics analysis. Volumes of interest (VOIs) were manually drawn layer by layer on different sequences including HR-VWMRI 3D T1WI and T1CE by the above two senior radiologists along the margin of the plaques, respectively. The images of the sample patient are shown in Figure 2.

Randomly selected 40 cases to evaluated interobserver and intraobserver reproducibility using intraclass correlation coefficients (ICC). Radiologist H manually outlined the VOIs twice within three months, ICCs>0.75 indicated good consistency of intraobserver. Radiologist M outlined the VOIs once, ICCs>0.75 indicated good consistency of interobserver. Radiologist H completed the remaining outline.

1 www.itk-snap.org

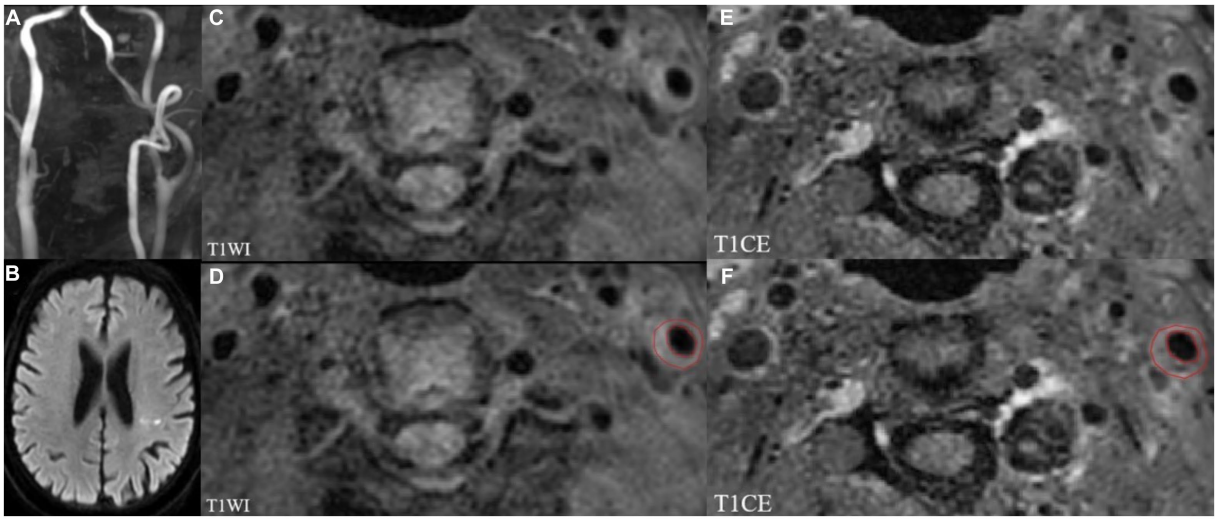


FIGURE 2
MRI images showing left internal carotid atherosclerotic plaque in a stroke patient. TOF-MRA **(A)** demonstrates mild stenosis, and DWI **(B)** shows the acute infarcts which are scattered and patchy in distribution within the left lateral ventricle posterior horn around. Left internal carotid atherosclerotic plaque on T1WI and T1CE images shown in panels **(C,E)**. The regions of interest (ROIs) of the carotid atherosclerotic plaque on T1WI and T1CE images **(D,F)**.

Feature extraction, selection, and model development

HR-VWMRI 3D T1WI and 3D T1CE images were used for feature extraction by applying the Pyradiomics package of Python software.[2] The extracted features include shape features, first-order statistics features, second-order statistics features, and higher order statistics features. In order to standardize and normalize the original features of different dimensions, the min–max standardization method is used to linearly transform the original feature dataset and map the values between 0 and 1. First, the features of $p < 0.05$ in each sequence were selected using Student's *t*-test. Second, the LASSO algorithm was used for selecting optimal feature subsets by penalty function λ adjust and set the corresponding coefficient of features with weak correlation to 0. Selection of the tuning parameter (λ) in the LASSO model via 10-fold cross-validation based on minimum criteria. The final retained features with non-zero coefficients and using non-zero coefficient features construct radiomics model. The radiomics score (Rad-Score) calculation used a linear combination of select parameters weighted by the relevant LASSO coefficients.

We selected the clinical and radiological characteristics with $p < 0.05$ in univariate analysis for multivariate logistic regression analysis, by which the odds ratios (ORs) with 95% confidence intervals (CIs) were calculated. The variables with $p < 0.05$ in multivariate analysis were finally used to establish the traditional model. The traditional model and radiomics features were combined to establish the combined model, which were displayed as radiomics nomograms (flowchart in Figure 3). In the training process of all models, to avoid overfitting of models, a 10-fold cross-validation method was used, and then, the prediction performance of all models in the training and test cohorts was evaluated, respectively.

Statistical analysis

All statistical analyses were performed using the IBM SPSS Statistics software (version 25.0) and Python software (see text footnote 2). The relationship between each variable and stroke status was evaluated by univariate analysis. The continuous variables adopted the Mann–Whitney *U*-test, and the categorical variables adopted the chi-squared test. Variables with $p < 0.05$ in univariate analysis were enrolled in multivariate logistic regression analysis. Receiver operating characteristic (ROC) analysis was used to determine the AUC values to evaluate the predictive performance of all models in both the training and test cohorts.

Results

Patient characteristics

In total, 127 patients with carotid plaque were enrolled in the final analysis, there were 60 stroke patients and 67 stroke-free patients. Table 2 lists the demographic and clinical characteristics of the enrolled patients.

Traditional model

Univariate analysis showed that gender, BMI, IPH, disrupted surface, enhancement, Max WA, NWI, and degree of stenosis were significantly associated with stroke (all $p < 0.05$, Table 2). Multivariate

2 https://www.python.org

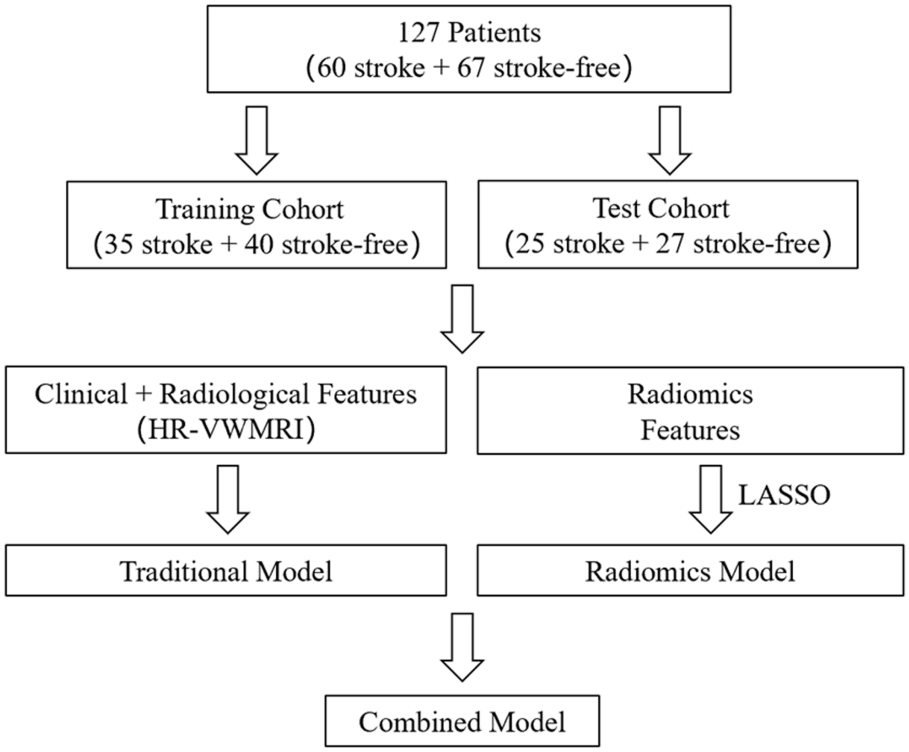


FIGURE 3
Flowchart for building the models predicting ischemic stroke.

TABLE 2 Demographic and clinical characteristics of the enrolled patients.

	n/total (%)	Stroke	Stroke-free	*p*-value[b]	Multivariate OR (95 %CI)[c]	*p*-value[c]	AUC
Sex	127	60	67	0.015[d]	0.168 (0.013, 2.205)	0.174	
Male	103	54	49				
Female	24	6	18				
Age[a]	60.55 ± 9.72	59.90 ± 10.90	61.20 ± 8.52	0.428[e]			
BMI[a]	24.81 ± 1.59	25.48 ± 1.68	24.22 ± 1.24	< 0.001[e]	2.564 (1.412, 4.656)	0.002	0.726
Hypertension	82	34	48	0.078[d]			
Hyperglycemia	43	22	21	0.527[d]			
Hyperlipidemia	36	15	21	0.428[d]			
Hyperuricemia	8	5	3	0.598[d]			
Hyperhomocysteinemia	75	39	36	0.197[d]			
IPH	45	36	9	< 0.001[d]	0.047 (0.005, 0.444)	0.008	0.733
LRNC	64	31	33	0.786[d]			
Disrupted surface	43	33	10	< 0.001[d]	0.527 (0.112, 2.490)	0.419	
Enhancement	41	25	16	0.017[d]	0.439 (0.082, 2.344)	0.335	
Remodeling pattern	127	60	67	0.509[d]			
Positive	80	36	44				
Negative	47	24	23				
Max WA (cm^2)[a]	0.50 ± 0.23	0.56 ± 0.25	0.45 ± 0.18	0.017[e]	1.103 (0.047, 25.918)	0.951	
NWI[a]	0.75 ± 0.19	0.83 ± 0.15	0.67 ± 0.18	< 0.001[e]	0 (0, 119.173)	0.156	
Degree of stenosis (%)[a]	71.00 ± 23.00	81.00 ± 17.00	62.00 ± 23.00	< 0.001[e]	1.155 (0.994, 1.341)	0.060	

AUC, area under the curve; CI, confidence intervals; BMI, body mass index; IPH, intraplaque hemorrhage; LRNC, lipid-rich necrotic core; Max WA, maximum wall area; NWI, normalized wall index; OR, odds ratio.
[a]Continuous variables shown with mean ± standard deviation (SD), others are categorical variables; [b]univariate analysis; [c]results from multivariate logistic analysis; [d]chi-squared test; [e]Mann-Whitney U-test.

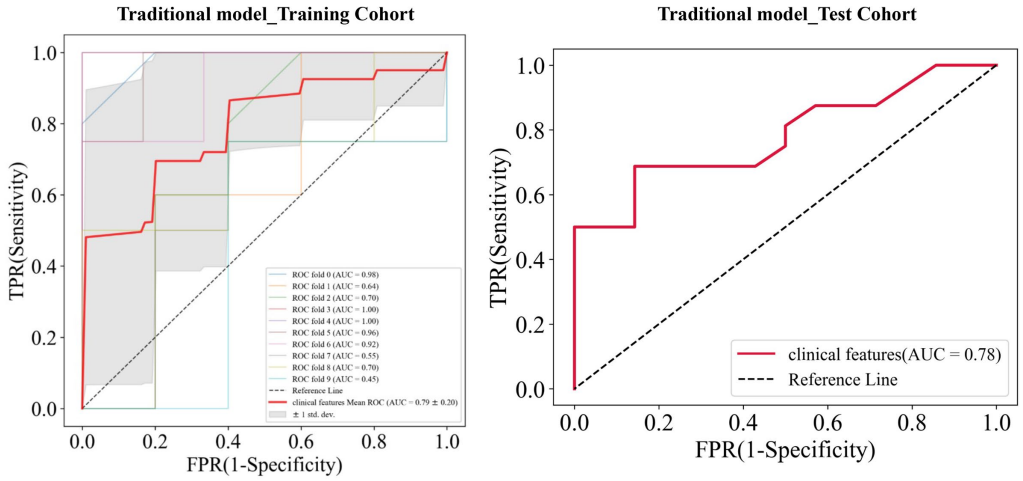


FIGURE 4
Receiver operating characteristic (ROC) curves of the traditional model in the training and test cohorts, respectively.

logistic regression analysis indicated the BMI (OR = 2.564; 95% CI, 1.412–4.656) and IPH (OR = 0.047; 95% CI, 0.005–0.444) were independent predictors of ischemic stroke and were used to establish the traditional model. When combining BMI and IPH, the AUC values were 0.79 and 0.78 in the training and test cohorts, respectively (Figure 4).

Radiomics model

A total of 239 and 191 features were extracted from the whole plaque region based on T1WI and T1CE sequences, respectively. After the LASSO algorithm was applied, 10 and 8 features were finally retained, which were used to establish the radiomics_T1WI model and the radiomics_T1CE model, respectively. The radiomics_T1WI model includes one feature of shape, two feature of first-order statistics, and seven features of texture (one gray-level co-occurrence matrix (GLCM) features, two gray-level dependence matrix (GLDM) features, two gray-level size-zone matrix (GLSZM) features, one neighborhood gray-tone difference matrix (NGTDM) features, and one gray-level run-length matrix (GLRLM) features). The radiomics_T1CE model includes two features of first-order statistics and six features of texture (three gray-level size-zone matrix (GLSZM) features, one gray-level run-length matrix (GLRLM) features, one neighborhood gray-tone difference matrix (NGTDM) features, and one gray-level dependence matrix (GLDM) features). The screening process and final screening characteristics are shown in Figures 5, 6.

The AUC value of the training cohort of the radiomics_T1WI model was 0.72, while the AUC value of the test cohort was 0.69. The AUC value of the training cohort of the radiomics_T1CE model was 0.82, while the AUC value of the test cohort was 0.74 (Figure 7).

Combined model

Finally, combined models were constructed and displayed as nomograms (Figure 8). In the training cohort, the combined_T1WI model exhibited an AUC value of 0.78, which was 0.81 in the test cohort. The combined_T1CE model exhibited an AUC value of 0.84 in the training cohort and 0.82 in the test cohort (Figure 9).

Table 3 lists the specificity, sensitivity, accuracy, AUC, and negative predictive value (NPV) and positive predictive value (PPV) of the traditional, radiomics, and combined models. The combined_T1CE model showed a higher AUC value than the other models.

Discussion

At present, the research on the relationship between carotid atherosclerotic plaque and ischemic stroke mainly focuses on the assessment of basic imaging characteristics of plaque or the degree of vascular stenosis. Previous studies have reported that the composition of atherosclerotic plaque is closely related to the occurrence of ischemic stroke (21). However, there are limitations in exploring the relationship between plaques and ischemic stroke based solely on their basic imaging features. First, clinical risk factors such as age, gender, hypertension, hyperglycemia, hyperlipidemia, obesity, smoking, and alcohol consumption are closely related to ischemic stroke. Second, the evaluation of the basic imaging features of plaques is subjective and qualitative, and the results are greatly influenced by the personal factors of the researchers. However, beyond this traditional evaluation method, carotid plaque imaging combined with artificial intelligence to accurately predict the risk of ischemic stroke is needed. Radiomics convert medical images to quantitative indicators through high-throughput extraction by data evaluation algorithms for predicting the risk of disease (22, 23). Therefore, carotid plaque VOIs were delineated on HR-VWMRI; traditional, radiomics, and combined models for predicting ischemic stroke were established.

Multivariate logistic regression analysis between stroke group and stroke-free group results showed that BMI and IPH were independent predictors of ischemic stroke. Then, the BMI and IPH were used to build the traditional model. The results revealed that the AUC of the traditional model is 0.79 in the training cohort and 0.78 in the test cohort. BMI is a commonly used indicator to measure the degree of

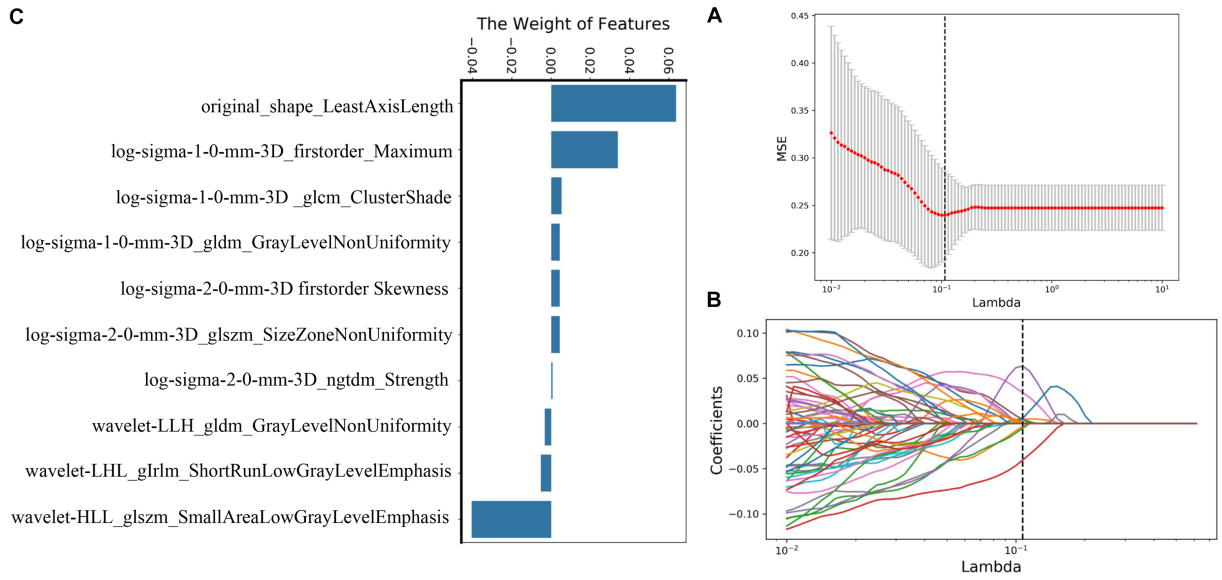


FIGURE 5
Selection of radiomics features using LASSO logistic regression based on T1WI images. **(A)** Selection of the tuning parameter (λ) in the LASSO model via 10-fold cross-validation based on minimum criteria. **(B)** The coefficients have been plotted versus (λ). **(C)** The final retained features with non-zero coefficients.

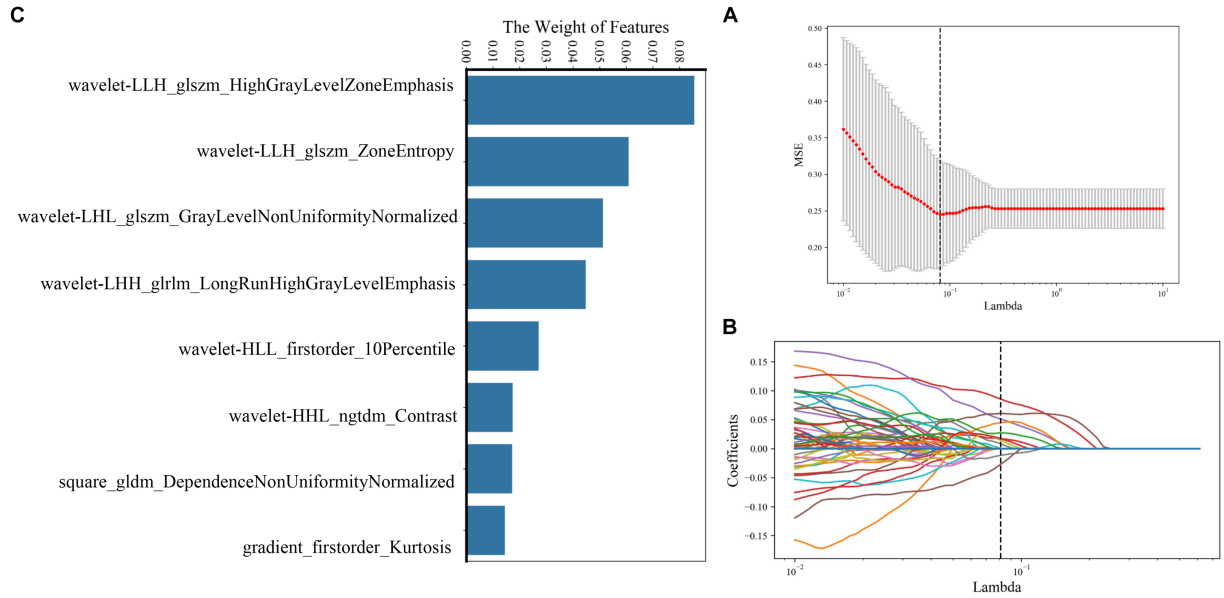


FIGURE 6
Selection of radiomics features using LASSO logistic regression based on T1CE images. **(A)** Selection of the tuning parameter (λ) in the LASSO model via 10-fold cross-validation based on minimum criteria. **(B)** The coefficients have been plotted versus (λ). **(C)** The final retained features with non-zero coefficients.

obesity and thinness of the body in the world, mainly used to reflect the total body fat (24). High BMI is closely related to hypertension, diabetes, and other risk factors, which are collectively called metabolic syndrome (25, 26). Studies found that the components of metabolic syndrome interacted with each other to promote the progress of metabolic disorder *in vivo*, which not only led to intracranial and extracranial atherosclerotic lesions but also led to the impairment of cerebrovascular regulation ability and microcirculation, further promoting the occurrence and development of cerebrovascular diseases dominated by ischemic stroke (27, 28). Atherosclerosis is known to be the main cause of ischemic stroke. Research shows that fat cells in patients with high BMI significantly increase, promote the release of inflammatory cytokines, and change the inflammatory state of the body (29, 30). The formation and development of atherosclerotic

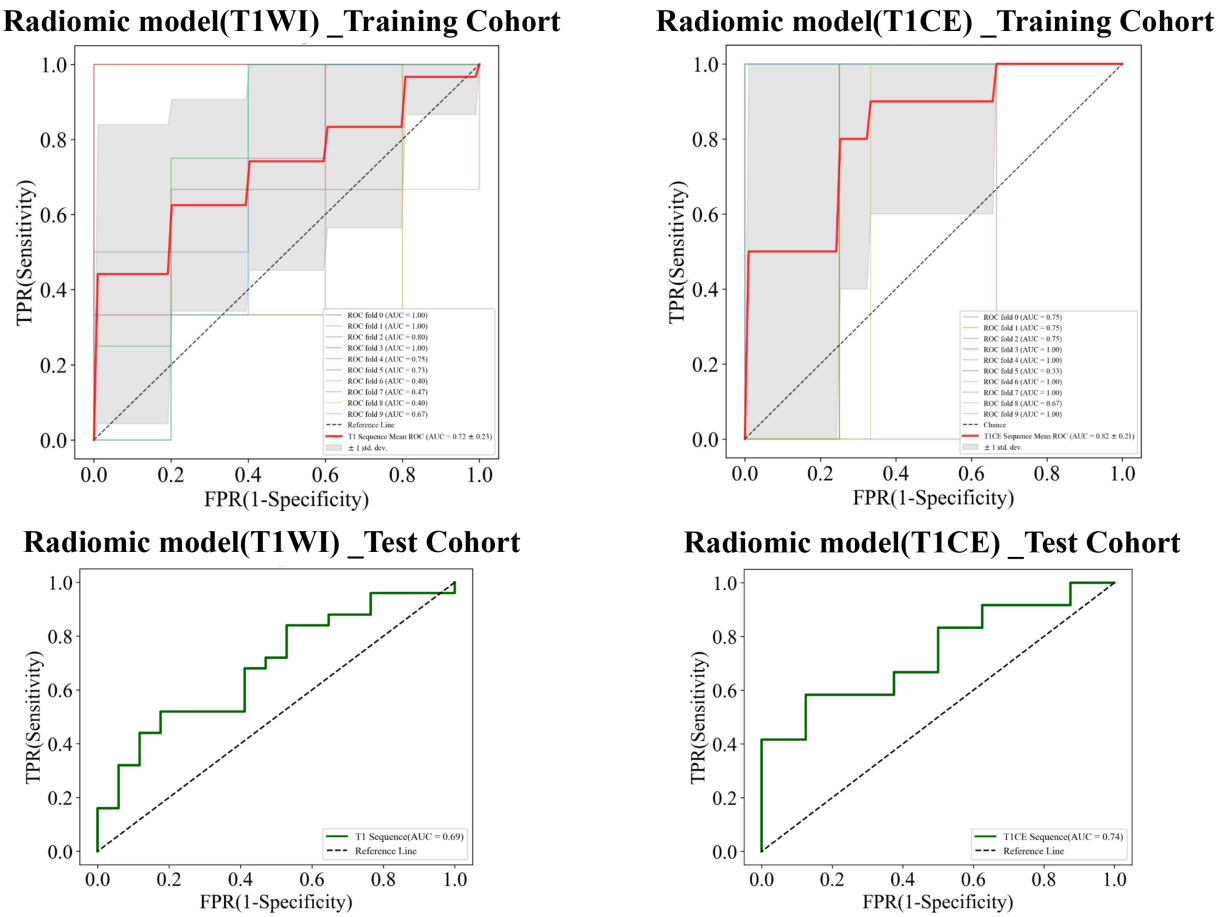


FIGURE 7
Receiver operating characteristic (ROC) curves of the radiomics_T1WI model and the radiomics_T1CE model in the training and test cohorts, respectively.

plaque is a chronic inflammatory process (31). Therefore, it is believed that high BMI is associated with ischemic stroke (32).

IPH is attributed to fragile neovascularization. The rupture of the neovascular endothelium will increase the stress of the plaque wall, making the plaque wall easier to rupture and cause thrombosis, which is more likely to lead to ischemic stroke (33, 34). Many studies have found that IPH is an independent predictor of stroke events. In patients with symptomatic atherosclerosis, the incidence of IPH is higher than that of asymptomatic patients (35–37). The results of our study revealed that the incidence of IPH in the stroke group was higher than that in the stroke-free group, which was consistent with previous study results. In the training and test cohorts, the AUC value of the radiomics_T1WI model is the lowest, but the prediction performance is significantly improved when the model combines IPH and BMI. In addition, when the radiomics_T1CE model is combined with IPH and BMI, the prediction performance of the model is further improved. This indicates that IPH and BMI are significantly associated with ischemic stroke (38).

The establishment of radiomics_T1WI and radiomics_T1CE models is based on HR-VWMRI 3D T1WI and 3D T1CE images, respectively. The use of HR-VWMRI 3D imaging can better characterize plaque features, including more comprehensive, rich, and detailed image information. 3D imaging methods can reduce local volume effects in 2D imaging and improve the results of radiomics analysis. We found the independent radiomics features from T1WI and T1CE images were different. This is because the signal characteristics in pre- and post-contrast T1WI reflect different pathophysiological characteristics of plaque. For example, the hyperintensity on pre-contrast T1WI is possibly IPH; on the other hand, the hyperintensity on post-contrast T1WI is attributed to the plaque neovascularization or the contrast uptake by active inflammation (39). After applying the LASSO algorithm, 10 features were finally retained based on T1WI including one shape feature, two first-order statistics features, and seven texture features, and eight features were finally retained based on T1CE including two first-order statistics features and six texture features. The first-order statistics features describe the distribution of single voxel value without considering the spatial relationship and are obtained based on histogram analysis and calculation (12). The second-order statistics features are usually described as "texture" features, which describe the statistical relationship between voxels with similar (or different) contrast values (11). In the 18 final features, only first-order statistics features and texture features appeared in the final screening results of the two sequences. It means that these two types of features may be the most important quantitative features to describe plaques, and these features cannot be visually evaluated by radiologists.

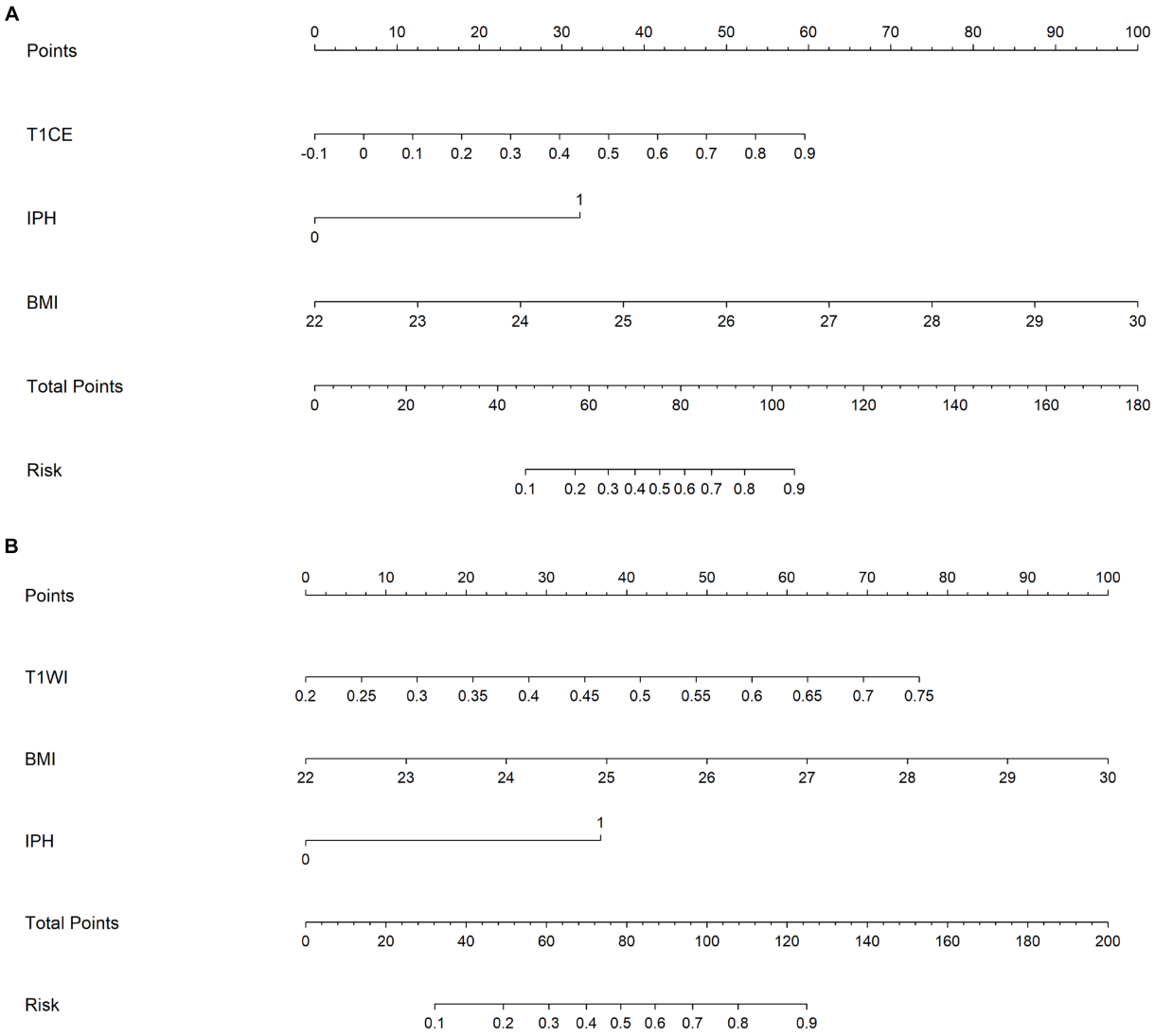


FIGURE 8
A nomogram which integrates the radiomics scores and traditional features of the training cohort. From the left to right, the probability scoring of ischemic stroke is marked on each axis and increases. An example of how to calculate the total points of a plaque on T1CE and predict the risk of the patient's ischemic stroke was as follows: a carotid plaque with IPH and the BMI of the patient is 25 receives 32.5 + 37.5 = 70 points from traditional features. A radiomics score of 0.7 corresponds to 47.5. Therefore, this patient scored 117.5 on the nomogram, which indicates a risk of ischemic stroke over 90%. **(A)** Combined_TICE model nomogram. **(B)** Combined_TIWI model nomogram.

Our study revealed that the prediction performance of the radiomics_T1CE model is significantly superior to the radiomics_T1WI model no matter in training or test cohort. This reflects that plaque enhancement is another independent risk factor for ischemic stroke (20, 40). The main reason for plaque enhancement is the increase in neovascularization and endothelial permeability. The contrast agent enters and stays in the plaque through the loose endothelium, resulting in plaque vulnerability and different degrees of enhancement. Therefore, plaque enhancement is closely related to the occurrence of ischemic stroke events (41, 42). However, it is interesting that the results of multivariate logistic regression analysis show that there is no statistically significant difference in plaque enhancement between the stroke group and the stroke-free group. This further shows that radiomics contains more information; for example, enhancement is a sign of high-risk plaques, but most previous studies were subjective visual qualitative recognition, lacking objective quantitative information. However, radiomics can provide quantitative information that is not relevant to the reader, which is difficult to visualize or too numerous for radiologists to visually evaluate (43).

In the training and test cohorts, the combined_T1CE model has the highest AUC value. Compared with the radiomics_T1CE model, the prediction performance of the combined_T1CE model has been improved. However, there is no statistically significant difference between the two models in the training cohort. This could be explained by the relative weights of the radiomics_T1CE model versus the traditional model, and the combined_T1CE model was weighted heavily toward the enhancement of radiomics characteristics, which produced better performances (10).

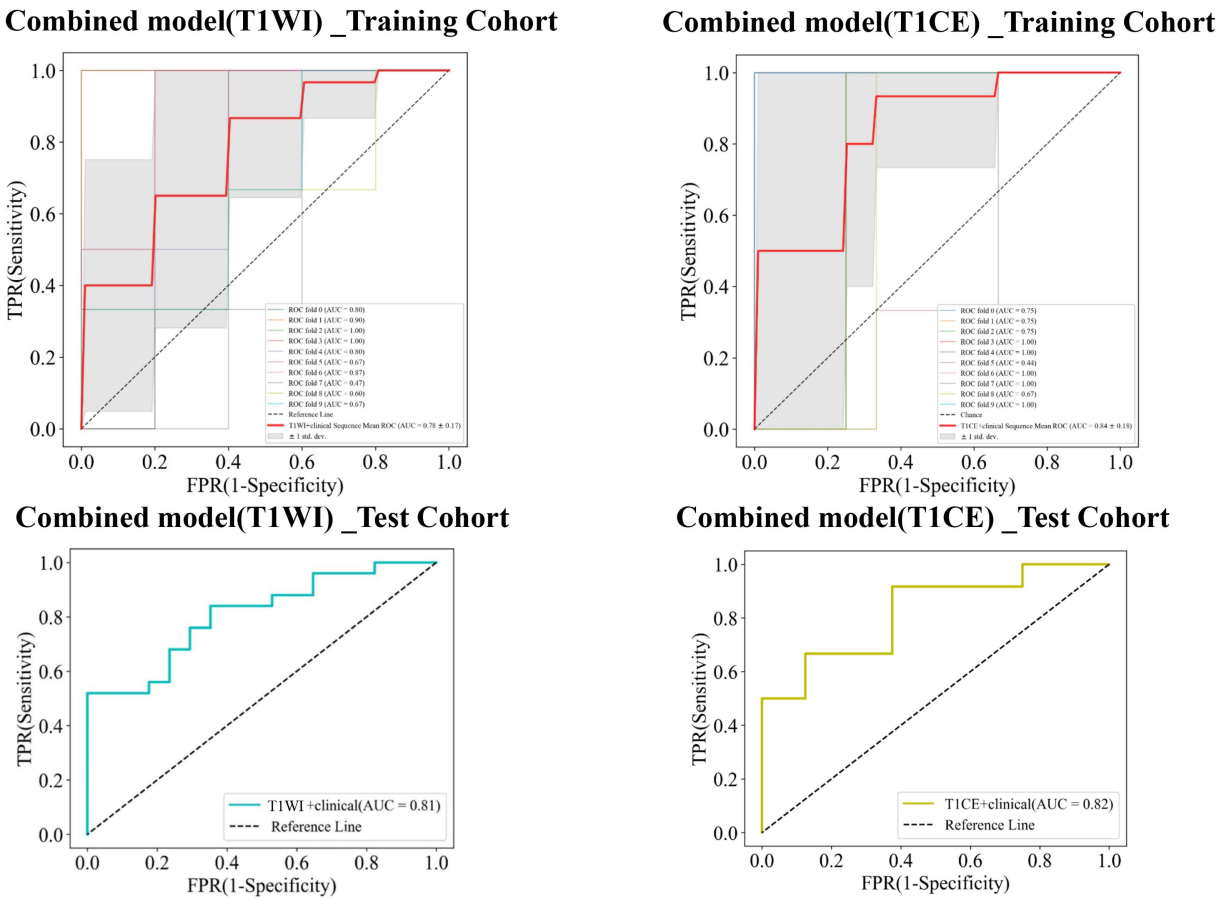


FIGURE 9
Receiver operating characteristic (ROC) curves of the combined_T1WI model and the combined_T1CE model in the training and test cohorts, respectively.

TABLE 3 **Predictive ability of all models.**

	Cohort	Sensitivity	Specificity	Accuracy	AUC	NPV	PPV
Traditional model	Training	0.61	0.78	0.70	0.79	0.73	0.69
	Test	0.69	0.85	0.77	0.78	0.70	0.85
Radiomics_T1WI model	Training	0.59	0.82	0.74	0.72	0.79	0.67
	Test	0.32	0.88	0.55	0.69	0.47	0.80
Radiomics_T1CE model	Training	0.72	0.82	0.77	0.82	0.80	0.77
	Test	0.50	0.88	0.65	0.74	0.54	0.86
Combined_T1WI model	Training	0.25	0.80	0.56	0.78	0.57	0.50
	Test	0.52	0.94	0.69	0.81	0.57	0.93
Combined_T1CE model	Training	0.82	0.78	0.79	0.84	0.86	0.79
	Test	0.50	0.88	0.70	0.82	0.57	0.84

This study beyond the assessment of the basic imaging characteristics of carotid atherosclerotic plaque or the degree of vascular stenosis uses a new model, that is, the combination of carotid atherosclerotic plaque imaging and artificial intelligence to explore the relationship with ischemic stroke. The study showed that radiomics score, IPH, and BMI were independent indicators of the risk of ischemic stroke. Combined with these independent risk factors, novel radiomics nomograms were generated. The generated nomogram based on T1CE had good predictive value, with AUCs of 0.84 and 0.82 in the training and test cohorts, respectively. These encouraging results deserve further multicenter trials applying carotid plaque radiomics features based on HR-VWMRI T1CE images and clinical characteristics for predicting the risk of ischemic stroke.

Limitations of this study: First, this study was a single center with a relatively small sample size; multicenter and larger data sets are needed to evaluate the prediction performance of the model in future study. Second, VOIs were manually delineated, despite the excellent reproducibility; however, due to the small size of the plaque, manual segmentation is a challenging task that takes a lot of time. Automatic segmentation would improve segmentation efficiency and accuracy. Third, because the boundary of plaque on TOF images is indistinct, and T2WI is not a 3D isovoxel sequence, we did not perform radiomics analysis on it. Fourth, there is a lack of multimodal imaging indicators such as cerebral blood flow and collateral circulation. In future, multimodality imaging combined with artificial intelligence is needed to establish a prediction model of ischemic stroke and improve the primary prevention strategy of stroke.

Conclusion

As a feasible and exploratory study, this study provides new insights into the prediction of ischemic stroke. The above results indicate that the combined_T1CE model incorporating clinical characteristics and carotid plaque radiomics features based on HR-VWMRI T1CE images can accurately predict the risk of ischemic stroke.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Review Board of Lanzhou University Second Hospital (scientific research project ethics approval number: 2022A-695). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NH: Writing – original draft, Methodology. WH: Writing – review & editing. YM: Methodology, Writing – review & editing. YZ: Data curation, Writing – review & editing. SY: Validation, Writing – review & editing. LM: Supervision, Writing – review & editing. JL: Formal analysis, Writing – review & editing. JZ: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the National Natural Science Foundation of China (grant number 81960309); Lanzhou University Second Hospital "Cuiying Technology Innovation Plan" Applied Basic Research Project (grant number CY2018-MS02); Gansu Province Clinical Research Center for Functional and Molecular Imaging (grant number 21JR7RA438); Special Fund for the Cultivation of Doctoral Candidates in Lanzhou University Second Hospital (grant number YJS-BD-28); Scientific Research Cultivation Plan of Cuiying Students in Lanzhou University Second Hospital (grant number CYXZ2021-33); and Lanzhou Talent Innovation and Entrepreneurship Project (grant number 2019-RC-95).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics–2015 update: a report from the American Heart Association. *Circulation*. (2015) 131:e29–e322. doi: 10.1161/CIR.0000000000000152

2. Sun T, Chen S, Wu K, Sun M, Zhang X, You C. Trends in incidence and mortality of stroke in China from 1990 to 2019. *Front Neurol*. (2021) 12:759221. doi: 10.3389/fneur.2021.759221

3. Ma Y, Cao J, Mubarik S, Bai J, Yang D, Zhao Y, et al. Age–period–cohort analysis of long trend of mortality for stroke and subtypes attributed to high SBP in Chinese adults. *Front Neurol*. (2022) 13:710744. doi: 10.3389/fneur.2022.710744

4. Xing L, Li R, Zhang S, Li D, Dong B, Zhou H, et al. High burden of carotid atherosclerosis in rural Northeast China: a population-based study. *Front Neurol*. (2021) 12:597992. doi: 10.3389/fneur.2021.597992

5. Saba L, Yuan C, Hatsukami TS, Balu N, Qiao Y, DeMarco J, et al. Carotid artery wall imaging: perspective and guidelines from the ASNR vessel wall imaging study group and expert consensus recommendations of the American Society of Neuroradiology. *Am J Neuroradiol*. (2018) 39:E9–E31. doi: 10.3174/ajnr.A5488

6. Yamada K, Yoshimura S, Shirakawa M, Uchida K, Maruyama F, Nakahara S, et al. High intensity signal in the plaque on routine 3D-TOF MRA is associated with ischemic stroke in the patients with low-grade carotid stenosis. *J Neurol Sci*. (2018) 385:164–7. doi: 10.1016/j.jns.2017.12.023

7. Lin G-H, Song JX, Fu NX, Huang X, Lu HX. Quantitative and qualitative analysis of atherosclerotic stenosis in the middle cerebral artery using high-resolution magnetic resonance imaging. *Can Assoc Radiol J*. (2021) 72:783–8. doi: 10.1177/0846537120961312

8. Meng Y, Mo Z, Hao J, Peng Y, Yan H, Mu J, et al. High-resolution intravascular magnetic resonance imaging of the coronary artery wall at 3.0 tesla: toward evaluation of atherosclerotic plaque vulnerability. *Quant Imaging Med Surg*. (2021) 11:4522–9. doi: 10.21037/qims-21-286

9. Saba L, Saam T, Jäger HR, Yuan C, Hatsukami TS, Saloner D, et al. Imaging biomarkers of vulnerable carotid plaques for stroke risk prediction and their potential clinical implications. *Lancet Neurol*. (2019) 18:559–72. doi: 10.1016/S1474-4422(19)30035-3

10. Zhang R, Zhang Q, Ji A, Lv P, Zhang J, Fu C, et al. Identification of high-risk carotid plaque with MRI-based radiomics and machine learning. *Eur Radiol*. (2021) 31:3116–26. doi: 10.1007/s00330-020-07361-z

11. Shi Z, Zhu C, Degnan AJ, Tian X, Li J, Chen L, et al. Identification of high-risk plaque features in intracranial atherosclerosis: initial experience using a radiomic approach. *Eur Radiol*. (2018) 28:3912–21. doi: 10.1007/s00330-018-5395-1

12. Shi Z, Li J, Zhao M, Peng W, Meddings Z, Jiang T, et al. Quantitative histogram analysis on intracranial atherosclerotic plaques: a high-resolution magnetic resonance imaging study. *Stroke*. (2020) 51:2161–9. doi: 10.1161/STROKEAHA.120.029062

13. Mitchell CC, Korcarz CE, Gepner AD, Nye R, Young RL, Matsuzaki M, et al. Carotid artery echolucency, texture features, and incident cardiovascular disease events: the MESA study. *J Am Heart Assoc*. (2019) 8:e010875. doi: 10.1161/JAHA.118.010875

14. Huang C, He Q, Huang M, Huang L, Zhao X, Yuan C, et al. Non-invasive identification of vulnerable atherosclerotic plaques using texture analysis in ultrasound carotid elastography: An in vivo feasibility study validated by magnetic resonance imaging. *Ultrasound in Medicine & Biology* (2017) 43:817–830.

15. Shi Z, Li J, Zhao M, Peng W, Meddings Z, Jiang T, et al. Quantitative histogram analysis on intracranial atherosclerotic plaques: a high-resolution magnetic resonance imaging study. *Stroke* (2020) 51:2161–2169.

16. Zhang R, Zhang Q, Ji A, Lv P, Zhang J, Fu C, et al. Identification of high-risk carotid plaque with MRI-based radiomics and machine learning. *Eur. Radiol.* (2021) 31:3116–3126.

17. Han N, Ma Y, Li Y, Zheng Y, Wu C, Gan T, et al. Imaging and hemodynamic characteristics of vulnerable carotid plaques and artificial intelligence applications in plaque classification and segmentation. *Brain Sci.* (2023) 13:143.

18. Lu M, Peng P, Cui Y, Qiao H, Li D, Cai J, et al. Association of progression of carotid artery wall volume and recurrent transient ischemic attack or stroke: a magnetic resonance imaging study. *Stroke*. (2018) 49:614–20. doi: 10.1161/STROKEAHA.117.019422

19. Cao X, Yang Q, Tang Y, Pan L, Lai M, Yu Z, et al. Normalized wall index, intraplaque hemorrhage and ulceration of carotid plaques correlate with the severity of ischemic stroke. *Atherosclerosis*. (2020) 315:138–44. doi: 10.1016/j.atherosclerosis.2020.10.896

20. Lee H, Ryu C, Yun SJ. Vessel-Wall magnetic resonance imaging of intracranial atherosclerotic plaque and ischemic stroke: a systematic review and meta-analysis. *Front Neurol*. (2018) 9:1032. doi: 10.3389/fneur.2018.01032

21. Zhang Y, Bai Y, Xie J, Wang J, He L, Huang M, et al. Carotid plaque components and other carotid artery features associated with risk of stroke: a systematic review and meta-analysis. *J Stroke Cerebrovasc Dis*. (2022) 31:106857. doi: 10.1016/j.jstrokecerebrovasdis.2022.106857

22. Wang H, Song B, Ye N, Ren J, Sun X, Dai Z, et al. Machine learning-based multiparametric MRI radiomics for predicting the aggressiveness of papillary thyroid carcinoma. *Eur J Radiol*. (2020) 122:108755. doi: 10.1016/j.ejrad.2019.108755

23. Ren J, Tian J, Yuan Y, Dong D, Li X, Shi Y, et al. Magnetic resonance imaging based radiomics signature for the preoperative discrimination of stage I-II and III-IV head and neck squamous cell carcinoma. *Eur J Radiol*. (2018) 106:1–6. doi: 10.1016/j.ejrad.2018.07.002

24. Fan J, Li X, Yu X, Liu Z, Jiang Y, Fang Y, et al. Global burden, risk factor analysis, and prediction study of ischemic stroke, 1990–2030. *Neurology*. (2023) 101:e137–50. doi: 10.1212/WNL.0000000000207387

25. Flegal KM. BMI and obesity trends in Chinese national survey data. *Lancet*. (2021) 398:5–7. doi: 10.1016/S0140-6736(21)00892-8

26. Marini S, Merino J, Montgomery BE, Malik R, Sudlow CL, Dichgans M, et al. Mendelian randomization study of obesity and cerebrovascular disease. *Ann Neurol*. (2020) 87:516–24. doi: 10.1002/ana.25686

27. Sarrafzadegan N, Gharipour M, Sadeghi M, Nezafati P, Talaie M, Oveisgharan S, et al. Metabolic syndrome and the risk of ischemic stroke. *J Stroke Cerebrovasc Dis*. (2017) 26:286–94. doi: 10.1016/j.jstrokecerebrovasdis.2016.09.019

28. Wang Z, Lu J, Weng W, Zhang L, Zhang J. Women's reproductive traits and ischemic stroke: a two-sample Mendelian randomization study. *Ann Clin Transl Neurol*. (2023) 10:70–83. doi: 10.1002/acn3.51702

29. Xiong Y, Zhang F, Zhang Y, Wang W, Ran Y, Wu C, et al. Insights into modifiable risk factors of erectile dysfunction, a wide-angled Mendelian randomization study. *J Adv Res*. (2023). doi: 10.1016/j.jare.2023.05.008, [Online ahead of print]

30. Zhang J, Wang A, Tian X, Meng X, Xie X, Jing J, et al. Impact of body mass index on efficacy and safety of ticagrelor versus clopidogrel in patients with minor stroke or transient ischemic attack. *CMAJ*. (2023) 195:E897–904. doi: 10.1503/cmaj.230262

31. Rohm TV, Meier DT, Olefsky JM, Donath MY. Inflammation in obesity, diabetes, and related disorders. *Immunity*. (2022) 55:31–55. doi: 10.1016/j.immuni.2021.12.013

32. Wang M, Zhang Z, Daghlas I, Gill D, Liu D, Lian X, et al. Adiposity and functional outcome after ischemic stroke: a Mendelian randomization study. *Neurology*. (2024) 102:e208080. doi: 10.1212/WNL.0000000000208080

33. McNally JS, McLaughlin MS, Hinckley PJ, Treiman SM, Stoddard GJ, Parker DL, et al. Intraluminal thrombus, intraplaque hemorrhage, plaque thickness, and current smoking optimally predict carotid stroke. *Stroke*. (2015) 46:84–90. doi: 10.1161/STROKEAHA.114.006286

34. Lu M, Zhang L, Yuan F, Peng P, Zhang H, Liu S, et al. Comparison of carotid atherosclerotic plaque characteristics between symptomatic patients with transient ischemic attack and stroke using high-resolution magnetic resonance imaging. *BMC Cardiovasc Disord*. (2022) 22:1–8. doi: 10.1186/s12872-022-02624-7

35. Larson A, Brinjikji W, Savastano L, Rabinstein AA, Saba L, Huston J, et al. Carotid intraplaque hemorrhage and stenosis: at what stage of plaque progression does intraplaque hemorrhage occur, and when is it most likely to be associated with symptoms? *Am J Neuroradiol*. (2021) 42:1285–90. doi: 10.3174/ajnr.A7133

36. Zhu C, Tian X, Degnan AJ, Shi Z, Zhang X, Chen L, et al. Clinical significance of intraplaque hemorrhage in low-and high-grade basilar artery stenosis on high-resolution MRI. *Am J Neuroradiol*. (2018) 39:1286–92. doi: 10.3174/ajnr.A5676

37. Zhou T, Jia S, Wang X, Wang B, Wang Z, Wu T, et al. Diagnostic performance of MRI for detecting intraplaque hemorrhage in the carotid arteries: a meta-analysis. *Eur Radiol*. (2019) 29:5129–38. doi: 10.1007/s00330-019-06053-7

38. Wei Y, Wang Z, He Q, Siddiqi SM, Zhou Z, Liu L, et al. Inverse association between plasma phylloquinone and risk of ischemic stroke in Chinese adults with hypertension and high BMI: a nested case-control study. *J Nutr*. (2022) 152:1927–35. doi: 10.1093/jn/nxac131

39. Yang W-J, Chen XY, Zhao HL, Niu CB, Zhang B, Xu Y, et al. Postmortem study of validation of low signal on fat-suppressed T1-weighted magnetic resonance imaging as marker of lipid core in middle cerebral artery atherosclerosis. *Stroke*. (2016) 47:2299–304. doi: 10.1161/STROKEAHA.116.013398

40. Gupta A, Baradaran H, al-Dasuqi K, Knight-Greenfield A, Giambrone AE, Delgado D, et al. Gadolinium enhancement in intracranial atherosclerotic plaque and ischemic stroke: a systematic review and meta-analysis. *J Am Heart Assoc*. (2016) 5:e003816. doi: 10.1161/JAHA.116.003816

41. Zhang X, Chen L, Li S, Shi Z, Tian X, Peng W, et al. Enhancement characteristics of middle cerebral arterial atherosclerotic plaques over time and their correlation with stroke recurrence. *J Magn Reson Imaging*. (2021) 53:953–62. doi: 10.1002/jmri.27351

42. Hou Z, Li M, Lyu J, Xu Z, Liu Y, He J, et al. Intraplaque enhancement is associated with artery-to-artery embolism in symptomatic vertebrobasilar atherosclerotic diseases. *Front Neurol*. (2021) 12:680827. doi: 10.3389/fneur.2021.680827

43. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. (2016) 278:563–77. doi: 10.1148/radiol.2015151169

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 19 March 2024
DOI 10.3389/fneur.2024.1330497

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Jawed Nawabi,
Charité University Medicine Berlin, Germany
Kersten Villringer,
Charité University Medicine Berlin, Germany

*CORRESPONDENCE
Helge C. Kniep
✉ h.kniep@uke.de

†These authors have contributed equally to this work and share first authorship

‡These authors have contributed equally to this work and share senior authorship

RECEIVED 30 October 2023
ACCEPTED 05 March 2024
PUBLISHED 19 March 2024

CITATION
Palsson F, Forkert ND, Meyer L, Broocks G, Flottmann F, Maros ME, Bechstein M, Winkelmeier L, Schlemm E, Fiehler J, Gellißen S and Kniep HC (2024) Prediction of tissue outcome in acute ischemic stroke based on single-phase CT angiography at admission.
Front. Neurol. 15:1330497.
doi: 10.3389/fneur.2024.1330497

COPYRIGHT
© 2024 Palsson, Forkert, Meyer, Broocks, Flottmann, Maros, Bechstein, Winkelmeier, Schlemm, Fiehler, Gellißen and Kniep. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Prediction of tissue outcome in acute ischemic stroke based on single-phase CT angiography at admission

Frosti Palsson[1,2†], Nils D. Forkert[3,4,5,6†], Lukas Meyer[2], Gabriel Broocks[2], Fabian Flottmann[2], Máté E. Maros[2], Matthias Bechstein[2], Laurens Winkelmeier[2], Eckhard Schlemm[7], Jens Fiehler[2], Susanne Gellißen[2‡] and Helge C. Kniep[2*‡]

[1]deCODE Genetics Inc., Reykjavik, Iceland, [2]Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, [3]Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, [4]Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, [5]Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, [6]Hotchkiss Brain Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, [7]Department of Neurology, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

Introduction: In acute ischemic stroke, prediction of the tissue outcome after reperfusion can be used to identify patients that might benefit from mechanical thrombectomy (MT). The aim of this work was to develop a deep learning model that can predict the follow-up infarct location and extent exclusively based on acute single-phase computed tomography angiography (CTA) datasets. In comparison to CT perfusion (CTP), CTA imaging is more widely available, less prone to artifacts, and the established standard of care in acute stroke imaging protocols. Furthermore, recent RCTs have shown that also patients with large established infarctions benefit from MT, which might not have been selected for MT based on CTP core/penumbra mismatch analysis.

Methods: All patients with acute large vessel occlusion of the anterior circulation treated at our institution between 12/2015 and 12/2020 were screened ($N = 404$) and 238 patients undergoing MT with successful reperfusion were included for final analysis. Ground truth infarct lesions were segmented on 24 h follow-up CT scans. Pre-processed CTA images were used as input for a U-Net-based convolutional neural network trained for lesion prediction, enhanced with a spatial and channel-wise squeeze-and-excitation block. Post-processing was applied to remove small predicted lesion components. The model was evaluated using a 5-fold cross-validation and a separate test set with Dice similarity coefficient (DSC) as the primary metric and average volume error as the secondary metric.

Results: The mean ± standard deviation test set DSC over all folds after post-processing was 0.35 ± 0.2 and the mean test set average volume error was 11.5 mL. The performance was relatively uniform across models with the best model according to the DSC achieved a score of 0.37 ± 0.2 after post-processing and the best model in terms of average volume error yielded 3.9 mL.

Conclusion: 24 h follow-up infarct prediction using acute CTA imaging exclusively is feasible with DSC measures comparable to results of CTP-based algorithms reported in other studies. The proposed method might pave the

way to a wider acceptance, feasibility, and applicability of follow-up infarct prediction based on artificial intelligence.

KEYWORDS

stroke, infarct core, mechanical thrombectomy, deep learning, segmentation

1 Introduction

Acute ischemic stroke is a leading cause of death and disability (1). Thrombolysis with recombinant tissue plasminogen activator (rtPA) had been the only treatment option for many years until multiple randomized controlled trials (2) confirmed high efficacy of mechanical thrombectomy (MT) in large vessel occlusions. Timely and accurate identification of the severity of the stroke with assessment of tissue infarction is critical for identifying patients that might benefit from MT. Recent advances in artificial intelligence (AI) and machine learning have led to the development of predictive models for stroke outcome using computed tomography (CT) image data and other imaging modalities (3–17). These models have shown promising results in predicting tissue infarction and the likelihood of treatment response, with potential implications for patient selection and timing of intervention (18, 19).

To date, proposed methods for tissue outcome prediction mainly utilize CT perfusion (CTP) datasets (18–20). However, accuracy of CTP-based tissue outcome prediction depends on the quality of the available CTP datasets. Factors that can influence the accuracy of CTP-based tissue outcome prediction include motion artifacts and problems related to deconvolution required to calculate the perfusion parameter maps (21). Moreover, CTP imaging is not available in all centers and not always considered a required standard of care in stroke imaging protocols.

Although perfusion imaging with assessment of core and penumbra has been established in many centers, growing evidence suggests that patient selection for MT using unenhanced CT and CTA only might also contribute to improved functional outcome. Recently published results of the TENSION RCT show that MT was associated with improved functional outcome and lower mortality in patients with established large infarct that were selected for MT based on non-contrast CT (22). These results are especially interesting as enrolled patients with large hypodense lesions at admission (ASPECTS 3–5) might not have been selected for endovascular therapy based on mismatch/perfusion although these patients benefit from MT.

Singe-phase CT angiography (CTA) is a non-invasive imaging modality that is considered standard of care in the evaluation of acute stroke patients. CTA imaging allows rapid identification of patients with large vessel occlusion (LVO) that may be amenable to endovascular treatment and provides imaging information of the collateral circulation, which is associated with patient prognosis (2). In comparison to CTP, CTA images can be acquired fast and with low technical effort and do not require costly and special licenses for acquisition and processing. Within this context, CTA datasets contain not only valuable information related to the clot location and collateral situation, but also on tissue edema formation visible as hypodense regions of the brain tissue. Despite the prognostic information available in CTA images, the value of CTA-based tissue outcome prediction using deep learning approaches has not been evaluated so far.

Thus, the goal of this work was to develop a deep learning-based algorithm that can predict the follow-up infarct location and volume based on single-phase CTA datasets only acquired acutely after patient admission. The proposed method is based on the well-established U-Net (23) architecture with several novel modifications such as residual blocks with spatial-, and channel-wise squeeze and excitation.

2 Materials and methods

2.1 Data availability

The data and code that support the findings of this study are available upon reasonable request from the corresponding author.

2.2 Study guidelines

The analysis was conducted in accordance with the "TRIPOD Checklist: Prediction Model Development and Validation."

2.3 Study population

The study was approved by the ethics committee of the chamber of physicians at Hamburg (MC-039/16), in accordance with the Declaration of Helsinki. All patients with anterior circulation stroke due to large vessel occlusion, age $\geq$ 18 years, and treated at our institution with endovascular procedure between December 2015 and December 2020 were retrospectively screened. For this analysis, all patients with anterior circulation stroke and successful recanalization defined as modified Thrombolysis in Cerebral Infarction (mTICI) Scale of 2b or 3 and availability of acute CTA imaging and 24 h follow-up non-contrast CT of the brain were included. Patients with failed recanalization were excluded to eliminate effects of persistent ischemia after MT.

2.4 Clinical and radiologic assessment

All clinical parameters including modified Rankin Scale (mRS), vessel occlusion status and location are site reported parameters (24, 25). Reperfusion success is assessed using the mTICI scoring system (26). mTICI scoring was conducted based on the initial occlusion location and the reperfusion success within the downstream territory of the initially occluded vessel/branch. Clinical assessments and reading of baseline imaging, digital subtraction angiograms and

follow-up imaging were conducted by local investigators at each participating center (single reader). Functional independence was defined as 90d mRS 0–2.

2.5 CT image acquisition

CT images at admission were acquired on a 2 × 128 slice scanner (SOMATOM Definition Flash, Siemens Healthcare GmbH, Erlangen, Germany) with the following imaging parameters: NCCT with 120 kV, 280 mA, less than 5.0 mm slice reconstruction and less than 0.5 mm in-plane; CTA: 100–120 kV, between 260 and 300 mA, 1.0 mm slice reconstruction, 0.5 mm collimation, 0.8 pitch, H20f soft kernel, 60 mL highly iodinated contrast medium and 30 mL NaCl flush at 4 mL/s; scan starts 6 s after bolus tracking at the level of the ascending aorta.

2.6 Infarct core segmentation

Manual segmentation of follow-up infarct volumes was performed using ITK-SNAP 3.8.0 (27) on 24 h follow-up CT scans (slice thickness 4.0 mm) by a senior neuroradiologist with more than 14 years of clinical experience (SG), blinded to clinical outcome data. In addition, segmentation results were visually verified by a second senior neuroradiologist with more than 20 years of clinical experience (JF). In case of disagreement, segmentations were reassessed by both readers and a consensus segmentation was generated.

2.7 Pre-processing

The segmentation approach uses a two-channel volume composed of the standard CTA at admission and a maximum intensity projection (MIP) as input. To eliminate uninformative rigid differences and to allow training based on the entire brain scans, input volumes were registered to a standard space CT brain atlas (28) resampled to 1x1x5 mm^3 voxel size.

In detail, the pre-processing procedure consisted of the following steps: (1) Automatic cropping of CTA datasets to the head region using the robustFOV FSL tool (29, 30); (2) Thresholding between 0 and 400 Hounsfield units; (3) Skull stripping; (4) Deriving 5 mm MIP reconstructions of the original CTA images; (5) Registration of the non-MIP images to the CT brain standard atlas (28), application of transformation to MIP-images, registration was performed using the antsRegistrationSyNQuick command from the ANTs toolbox (31) with rigid and affine transformations; (6) Registration of the scull-stripped 24 h FU CT scan to the CT brain standard atlas using the same toolbox with rigid and affine transformations, transformation of the ground truth infarct lesion labels to standard space.

All registration outcomes were visually verified. Cases were excluded if registration failed or image quality was inadequate for evaluation.

After pre-processing, the convolutional neural network (CNN) input volume was generated by combining the registered non-MIP and MIP images into a two-channel 4D volume. The final size of this volume was 192 × 224 × 32 × 2 voxels, with spatial dimensions divisible by 24, in line with the U-Net model's encoder stages described subsequently.

2.8 Model architecture

Figure 1 shows the architecture of our proposed lesion prediction model. It uses the well-established U-Net framework (23) as the basis, featuring a distinct encoder-decoder design with interspersed skip connections. The encoder handles feature extraction, while the decoder translates these extracted features back to the image domain. The architecture spans four stages, with each stage comprising a residual block (32). Notably, this block integrates two convolutional layers followed by a spatial and channel-wise squeeze and excite block (33) (SE block). The SE block is composed of parallel branches of spatial-and channel-wise SE blocks as shown in the lower right corner of Figure 1. The spatial part (upper branch) modulates the input feature map employing a learnable $1 \times 1 \times 1$ convolutional layer followed by a sigmoid activation to ensure output values are scaled between [0, 1]. The weights are arranged such that each output spatial element is a linear combination of all the different channels of the input feature map for this spatial location. Therefore, the output feature map is a spatially weighted version of the input feature map, where these weights are learned during training.

The channel-wise branch modulates the weights of the separate channels of the input feature map by passing it through a global average pooling layer followed by a dense layer with an output dimension that is half the number of channels, followed by ReLU activation and another dense layer with a dimension equal to the original number of channels. Finally, a sigmoid layer ensures that the output is scaled between [0, 1]. This weight tensor is then multiplied with the input feature map to obtain an output feature map with the individual channels scaled by the learned weights. The final output of the SE block is the sum of the outputs of the spatial- and channel-wise branches.

2.9 Model parameters

In medical image segmentation, the region of interest is often only a small part of the total area or volume, which is also the case for tissue outcome prediction tasks. This means that the ratio of foreground voxels to background voxels can be very small and, therefore, may lead to an imbalanced problem. To address this, we used a loss function based on the Dice similarity coefficient (DSC) (34), which is a popular choice for semantic segmentation as it enables efficient training despite class imbalance.

Another important function in neural network models is the activation function, which can impact the performance and training dynamics of the model. In this work, the recently proposed Mish activation function (35) was used, which is a self-regulating, non-monotonic activation function that has been shown to improve performance compared to other popular choices such as Rectified Linear Units (ReLU).

All models were trained for a total of 300 epochs and the validation DSC was used to select the best performing model based on a check-point system. The rectified ADAM optimizer (36) with an initial learning rate of 0.001 and exponential decay with the learning rate decaying by a factor of 0.25 every 50 epochs was used.

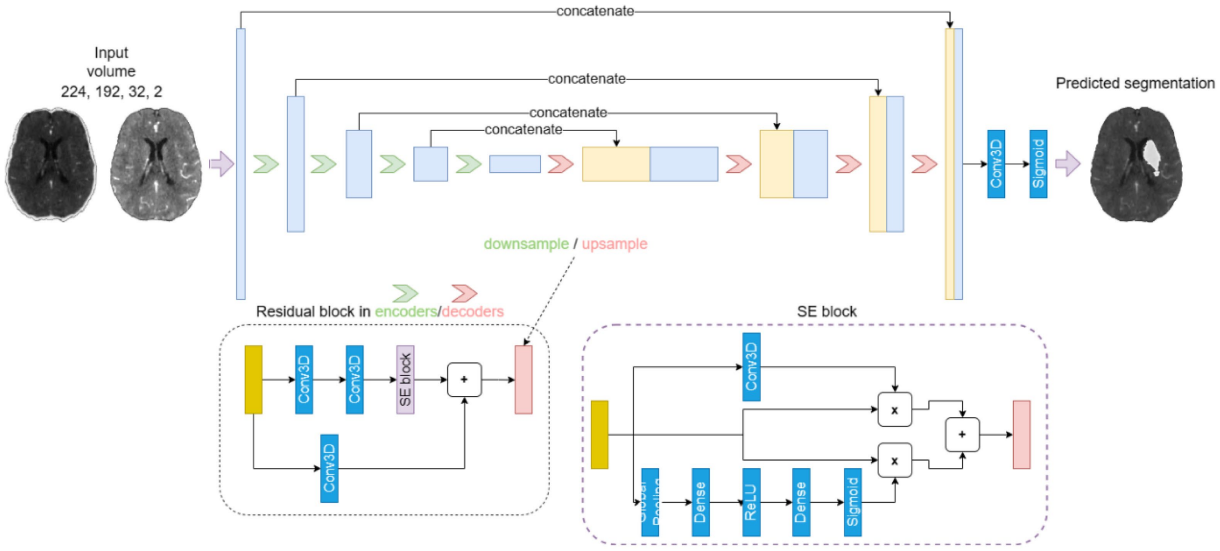


FIGURE 1
The U-Net-based 4-stage model architecture. Each stage is composed of a residual block (dotted box), which integrates two convolutional layers followed by a spatial and channel-wise squeeze and excite block (SE block, dashed box).

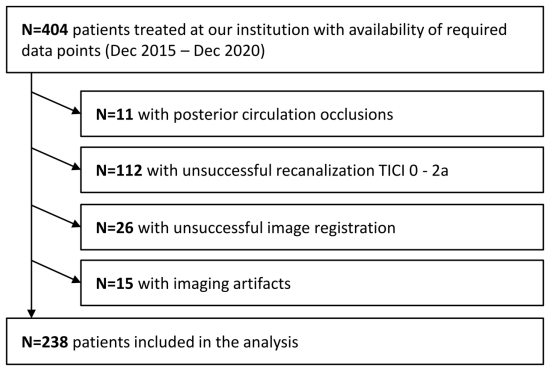


FIGURE 2
Patient inclusion flowchart.

2.10 Post-processing

To improve accuracy of the predicted lesion segmentation, we included a post-processing step that retained only the largest connected component of the predicted lesion. This decision is in line with the observation that most ground truth lesions consist of a single connected component. Since our model trains on entire images, its predictions predominantly reflect this characteristic. Nonetheless, occasionally, the model introduced secondary components in its predictions, which were usually inaccurate. By focusing on the largest connected component, we eliminated these erroneous predictions.

2.11 Experiment setup and evaluation metrics

A 5-fold nested cross-validation (CV) scheme was used for training, validation (i.e., hyper-parameter optimization) and testing. The available data ($n = 238$) was randomly split into five training/validation sets (80%, $n = 191$) and five separate test sets (20%, $n = 47$). Each training/validation set was again randomly split into 5 training sets (80% of each outer CV training/validation set, $n = 152$) and 5 validation sets (20% of each outer CV training/validation set, $n = 39$) for hyperparameter tuning. Random splits were conducted using a stratified approach based on the lesion size. This ensures a similar lesion size distribution in both sets. For each outer CV run (models 1 to 5) the average DSC of the corresponding test set was used as the primary evaluation metric while the average absolute lesion volume difference of the corresponding test set was used as a secondary evaluation metric.

2.12 Multivariable regression analysis

The association of segmented vs. predicted volumes with functional independence (90 day mRS 0–2) was analyzed using multivariable logistic regression. For initial neurological status (NIHSS at admission), a linearized association was assumed. Regression models were adjusted for age and pre-stroke mRS, adjusted odds ratios (aOR), coefficients, 95% confidence intervals and p-values were reported. p-values <0.05 were defined as statistically significant. Regression analysis was conducted with Stata/MP 18.0.

3 Results

A total of 404 patients were screened, and 238 patients were included in the analysis (Figure 2). Included patients had a median age of 76 years (IQR: 64; 81), a median NIHSS at admission of 16 (IQR: 11; 19), median ASPECTS of 7 (IQR: 6; 9), median infarct volume of 25 mL (IQR: 8; 114) and median 90-days mRS of 4 (IQR: 1; 5; Table 1). Table 2 displays the test performance of the five cross-validation models, including DSC and volume error (ml) for each outer fold test

TABLE 1 Baseline clinical characteristics of study cohort.

Variable	Median | n (%)	Q1; Q3	Range
Age (median)	76	64; 81	29–97
Sex (f)	121 (51%)		
Pre-stroke mRS (median)	0	0; 1	0–5
NIHSS at admission (median)	16	11; 19	0–42
Comorbidity hypertonus	158 (66%)		
Comorbidity diabetes	39 (16%)		
Comorbidity dyslipidemia	33 (14%)		
Comorbidity atrial fibrillation	88 (37%)		
ASPECTS at admission (median)	7	6; 9	1–10
i.v. thrombolysis	135 (57%)		
# of passes (median)	2	1; 2	0–8
AE vasospasm	5 (2%)		
AE clot migration/embolization	6 (3%)		
AE dissection/perforation	3 (1%)		
AE ICH	3 (1%)		
Final TICI			
-2b	115 (48%)		
−3	123 (52%)		
Follow-up infarct volume 24 h CT (ml)	24.8	8.3; 114.2	0–516.1
90-days mRS (median)	4	1; 5	0–6

AE, Adverse event; ASPECTS, Alberta Stroke Program Early CT Score; ICH, Intracranial hemorrhage; mRS, modified Rankin Scale; NIHSS, National Institute of Health Stroke Scale; Q1, 1st quartile; Q3, 3rd quartile; TICI, Thrombolysis in cerebral infarction.

TABLE 2 Test set results as mean ± standard deviation for each outer loop of the 5-fold cross-validation models without and with post-processing.

Without post-processing			With post-processing	
Model	DSC	Volume Error [ml]	DSC	Volume Error [ml]
1	0.36 ± 0.2	11.4 ± 79	0.37 ± 0.2	6.2 ± 79
2	0.33 ± 0.2	13.5 ± 83	0.33 ± 0.2	8.4 ± 83
3	0.34 ± 0.2	29.6 ± 78	0.34 ± 0.2	22.6 ± 78
4	0.34 ± 0.2	9.9 ± 78	0.34 ± 0.2	3.9 ± 80
5	0.35 ± 0.2	25.1 ± 77	0.36 ± 0.2	16.3 ± 77
Mean	0.34 ± 0.2	17.9 ± 79	0.35 ± 0.2	11.5 ± 79.4

DSC, Dice similarity coefficient.

set. The results are presented without and with post-processing. Figure 3 shows Bland–Altman plots comparing the true vs. the predicted lesion volumes in ml for each model.

Results suggest that all models perform similarly with rather small quantitative differences regarding the test set metrics with average DSC values ranging between 0.33 and 0.37 (SD: ±0.20). Likewise, the standard deviation of the DSC is similar for all models. Post-processing only slightly improved the average DSC but considerably reduced the average volume error (11.5 mL (SD: ±79.4 mL) after post-processing).

Bland–Altman plots in Figure 3 revealed that all models tend to overestimate the true infarct volume, which was also supported by the volume error (2). Furthermore, data indicated a trend for underestimating the volume of very large infarcts.

Figure 4 shows a visualization of the results obtained using the model with the highest average DSC compared to the ground truth for four exemplified patients. In these examples, the lesion was correctly located in all cases. However, the shape of predictions and ground truth do not fully align. It can be also seen that post-processing was primarily helpful to remove false lesions, especially those located in the contralateral hemisphere.

Table 3 shows the association of segmented in comparison to predicted volumes with initial neurological status (NIHSS at admission) and functional independence (90 day mRS 0–2). Results suggest that higher follow-up infarct volumes are associated with lower probability of functional independence at day 90 for segmented volumes (aOR: 0.98 [95% CI, 0.97–1.00], $p < 0.05$) and predicted volumes (aOR: 0.98 [95% CI, 0.96–0.99], $p < 0.01$). Both coefficients were not statistically significantly different. For NIHSS at admission a correlation with the predicted volume (Coeff: 0.07 [95% CI, 0.04–0.09], $p < 0.001$) was observed, however, for segmented volumes, the coefficient was not statistically significant ($p = 0.112$).

4 Discussion

In this analysis, we sought to evaluate if CTA datasets can be used for tissue outcome prediction in patients with ischemic stroke. The most important finding of this work is that a prediction of tissue outcome using exclusively CTA datasets is generally feasible using

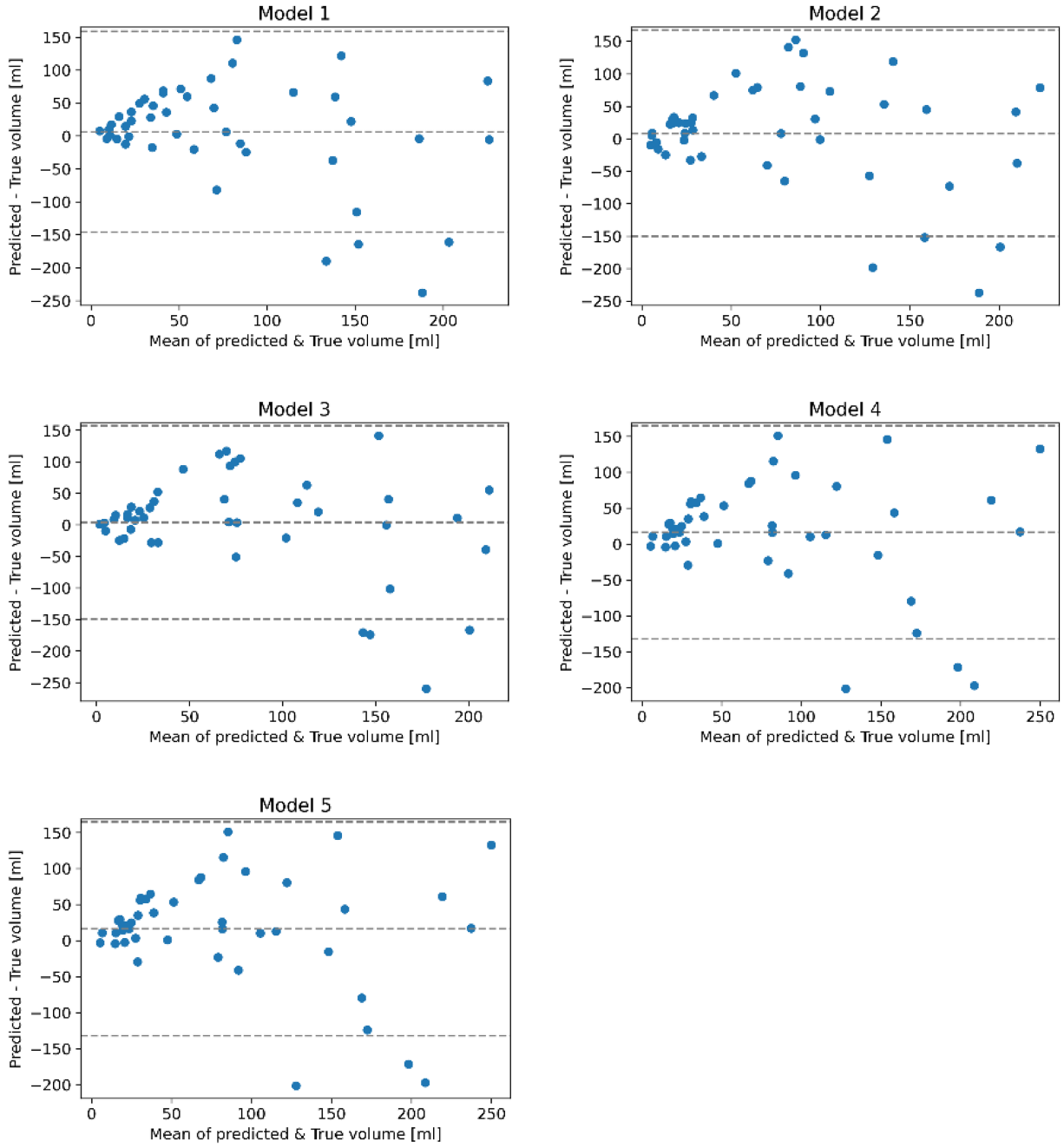


FIGURE 3
Bland–Altman plots comparing the true and predicted lesion volumes in ml for trained models on each loop of the 5-fold cross-validation after applying post-processing. Each plot is a scatterplot of the difference of the true and predicted volume vs. the mean of the predicted and true volume. The mean bias and regression based 95% limits of agreement are shown using dotted horizontal lines.

deep convolutional neural networks. Test set performance of our models reached an average DSC of 0.35 with mean volume error of 11.5 mL.

Overall, the predictive performance achieved in this work is within the range of previously described tissue outcome prediction models using more complex 4D CT perfusion datasets. For example, Amador et al. achieved an average DSC of 0.45 using a very advanced temporal convolutional neural network and 4D CTP datasets as input (20), while Qiu et al. achieved mean volume error of 21.7 mL using multiphase CTA images (4). It might be argued that a deep learning model using a single time-point CTA cannot outperform an advanced method having access to the complete hemodynamic perfusion information from a 4D CTP scan. However, CTP datasets typically used for lesion outcome prediction are often highly curated by excluding datasets with severe motion or other artifacts, which they are very sensitive to. This is one of the main benefits of the here proposed method, which uses simple single time-point CTA images that are less prone to motion and other artifacts and are widely available without any costly licenses for acquisition and processing.

Another benefit of the proposed method is that the input images are used en bloc in the model instead of splitting them into smaller patches during training and inference. In this way, the network sees the whole image at once, which might result in better learning of typical infarct locations and lesion distributions with regards to the entire brain. Within this context, better lesion outcome predictions may be possible with multi-phase CTA images or by combining CTA and CTP images (5). However, this would also require a more complex deep learning model that is capable of making use of the temporal information available in multi-phase CTA images and CTP images. Also, this approach would be more prone to movement artifacts and require more complex data leading to decreased feasibility and generalizability.

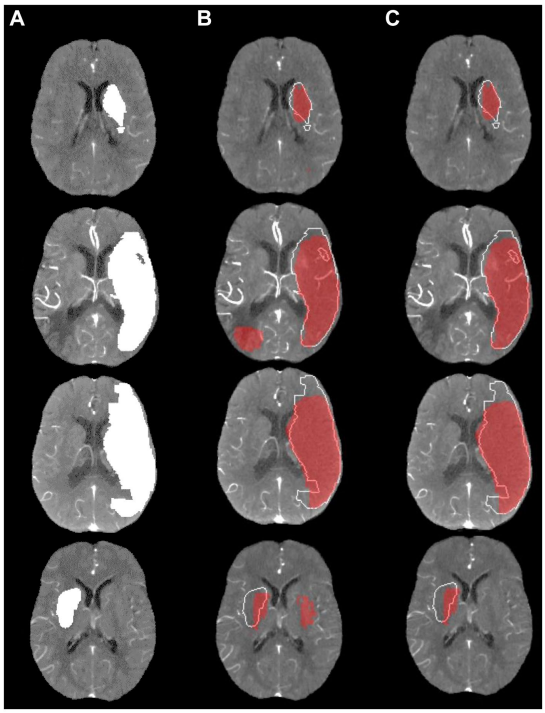


FIGURE 4
Visualization of lesion outcome prediction results obtained using the best model in terms of the average test set dice similarity coefficient for four exemplified patients. **(A)** Ground truth; **(B)** Prediction including small components red, ground truth white; **(C)** Prediction excluding small components red, ground truth white.

The quantitative results show that the proposed model leads to an underestimation of large lesions and overestimation of small lesions. This is a common problem of many segmentation methods (regression to the mean (37)). A potential solution to this problem may be a modification of the loss function to also include the volume error or to train multiple models for different lesion sizes (6). However, the second option would also reduce the number of datasets available for training of the lesion prediction model.

Furthermore, an overestimation of infarct volume in CTA source images has also been noted in previous studies (7) where it is postulated that a possible reason for this is that modern rapid-acquisition CT scanners may produce CTA images that are more strongly CBF- than CBV weighted (38), and therefore overestimate the true infarct volume. For example, in a study of 105 patients (8), it was found that follow-up infarct volume predictions based on CTA source images significantly overestimated the infarct size in many cases.

In line with results from previous studies, multivariable regression analysis suggest that follow-up infarct volume is associated with functional outcome (39–41). Furthermore, no statistically significant difference in association with functional outcome was observed for ground-truth follow-up volume segmentations and model predictions, suggesting that predicted infarct volumes might serve as additional surrogate marker for functional outcome. For NIHSS at admission a significant association with predicted infarct volumes was observed, however, segmented volumes were not significantly associated with NIHSS at admission. One explanation could be that predicted volumes are derived from CTA imaging at admission and might therefore better reflect neurological status at admission.

This work has multiple limitations that should be discussed. First, no comparison to CTP-based lesion outcome prediction methods was conducted as corresponding CTP datasets were not available for all

TABLE 3 (A) Multivariable logistic regression with functional independence (mRS 0–2 at 90 days) as dependent variable; (B) Multivariable linear regression with NIHSS at admission as dependent variable.

A: 90 days mRS 0–2 (Multivariable logistic regression)								
	Segmented volumes				Predicted volumes			
	aOR	*p* value	95% Conf. interval		aOR	*p* value	95% Conf. interval	
Volume (ml)	0.98	0.013	0.97	1.00	0.98	0.003	0.96	0.99
Age (years)	0.94	0.064	0.89	1.00	0.94	0.059	0.88	1.00
Pre-stroke mRS	0.82	0.681	0.33	2.07	0.96	0.939	0.36	2.54
Constant	109.89	0.025	1.78	6,780	376.34	0.011	3.87	36,605
B: NIHSS admission (Multivariable linear regression)								
	Segmented volumes				Predicted volumes			
	Coeff	P>t	95% Conf. interval		Coeff	p-value	95% Conf. interval	
Volume (ml)	0.02	0.112	0.00	0.04	0.07	<0.001	0.04	0.09
Age (years)	−0.07	0.373	−0.21	0.08	−0.06	0.321	−0.18	0.06
Pre-stroke mRS	0.13	0.906	−2.08	2.34	−0.78	0.398	−2.63	1.07
Constant	18.97	<0.001	9.33	28.62	14.93	0.001	6.85	23.01

mRS, Modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; aOR, Adjusted odds ratio; Coeff, Coefficient; CI, Confidence interval.

patients. Also, no comparison to non-enhanced CT-based methods was conducted. However, it can be assumed that tissue density information from non-enhanced CT is also included in CTA scans. Second, the proposed method did not include any clinical parameters in the model although they may improve the prediction accuracy. To date, there is no consensus regarding the optimal way to integrate clinical information in segmentation methods. Therefore, we restricted the proposed model to imaging information only to test the general feasibility of using only CTA datasets for this purpose. Third, the proposed method was only trained and tested using datasets from patients with a large vessel occlusion of the anterior circulation treated with MT. Thus, it remains unclear how well the proposed method can predict lesion outcomes in patients with an occlusion in the posterior circulation. Fourth, ground truth segmentation of the infarct core was conducted manually based on 24 h follow-up CT images. Interrater variability of manual segmentation processes might reduce the generalizability of results. However, the conducted visual verification of segmentations by a second reader and reassessment in case of disagreement increased validity of segmentation results. Fifth, model training and testing was conducted based on single-center data. However, besides external testing, the conducted nested cross validation is currently considered as gold standard for machine learning approaches. Although model training and testing was conducted based on single center data, it can be argued that CT imaging data of current state-of-the-art CT scanners has a high degree of standardization. Even if a certain center-specific bias cannot be excluded, it needs to be discussed if center-specific training and models on the other hand might allow higher predictive performance. In fact, several FDA-approved machine learning-based tools for acute stroke diagnostics require center-specific training of their algorithms. Sixth, the CNN was trained using 4D volumes generated from CTA images and MIP reconstruction resampled to 1 mm x 1 mm x 5 mm voxel size. Resampling to 5 mm slice thickness might reduce predictive precision especially for small infarcts.

5 Conclusion

24 h follow-up infarct prediction exclusively using acute single-phase CTA datasets is feasible and can be successfully achieved with good accuracy. In comparison to CTP data, CTA data is technically more widely available, in general incorporated into standard basic acute stroke protocols in clinical routine practice, and less prone to movement artifacts. The method proposed in this work based on single-phase CTA might pave the way to a wider acceptance, feasibility, and applicability of follow-up infarct prediction employing artificial intelligence methods.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics committee of the chamber of physicians at Hamburg. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

FP: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft. NF: Conceptualization, Data curation, Formal analysis, Supervision, Validation, Visualization, Writing – original draft. LM: Investigation, Project administration, Validation, Writing – review & editing. GB: Formal analysis, Investigation, Supervision, Validation, Writing – review & editing. FF: Conceptualization, Data curation, Project administration, Supervision, Validation, Writing – review & editing. MM: Project administration, Supervision, Validation, Writing – review & editing. MB: Conceptualization, Project administration, Supervision, Validation, Writing – review & editing. LW: Software, Supervision, Validation, Writing – review & editing. ES: Project administration, Supervision, Validation, Writing – review & editing. JF: Methodology, Project administration, Supervision, Validation, Writing – review & editing. SG: Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Validation, Writing – original draft. HK: Conceptualization, Methodology, Project administration, Supervision, Validation, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

HK and FF are consultants for Eppdata GmbH. HK has financial interest in Eppdata GmbH. NF has financial interest in Eppdata GmbH. JF is consultant for Cerenovus, Medtronic, Microvention, Penumbra, Phenox, Roche and Tonbridge, serves in the advisory board of Stryker and Phenox, is stock holder of Tegus Medical, Eppdata and Vastrax., and serves as Associate Editor at JNIS. Author FP was employed by company deCODE Genetics Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Donnan GA, Fisher M, Macleod M, Davis SM. Stroke. *Lancet*. (2008) 371:1612–23. doi: 10.1016/S0140-6736(08)60694-7

2. Goyal M, Menon BK, van Zwam WH, Dippel DW, Mitchell PJ, Demchuk AM, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. (2016) 387:1723–31. doi: 10.1016/S0140-6736(16)00163-X

3. Yedavalli VS, Tong E, Martin D, Yeom KW, Forkert ND. Artificial intelligence in stroke imaging: current and future perspectives. *Clin Imaging*. (2020) 69:246–54. doi: 10.1016/j.clinimag.2020.09.005

4. Qiu W, Kuang H, Ospel JM, Hill MD, Demchuk AM, Goyal M, et al. Automated prediction of ischemic brain tissue fate from multiphase computed tomographic angiography in patients with acute ischemic stroke using machine learning. *J stroke*. (2021) 23:234–43. doi: 10.5853/jos.2020.05064

5. McDougall CC, Chan L, Sachan S, Guo J, Sah RG, Menon BK, et al. Dynamic CTA-derived perfusion maps predict final infarct volume: the simple perfusion reconstruction algorithm. *AJNR Am J Neuroradiol*. (2020) 41:2034–40. doi: 10.3174/ajnr.A6783

6. Barros RS, Tolhuisen ML, Boers AM, Jansen I, Ponomareva E, Dippel DW. Automatic segmentation of cerebral infarcts in follow-up computed tomography images with convolutional neural networks. *J neurointerventional surgery*. (2020) 12:848–52. doi: 10.1136/neurintsurg-2019-015471

7. Hokkinen L, Mäkelä T, Savolainen S, Kangasniemi M. Evaluation of a CTA-based convolutional neural network for infarct volume prediction in anterior cerebral circulation ischaemic stroke. *European Radiology Experimental*. (2021) 5:25. doi: 10.1186/s41747-021-00225-1

8. Mukherjee A, Muthusami P, Mohimen A, Srinivasan K, Babunath B, Sylaja PN, et al. Noncontrast computed tomography versus computed tomography angiography source images for predicting final infarct size in anterior circulation acute ischemic stroke: a prospective cohort study. *J Stroke Cerebrovasc Dis*. (2017) 26:339–46. doi: 10.1016/j.jstrokecerebrovasdis.2016.09.026

9. Kidwell CS, Wintermark M, De Silva DA, Schaewe TJ, Jahan R, Starkman S, et al. Multiparametric MRI and CT models of infarct core and favorable penumbral imaging patterns in acute ischemic stroke. *Stroke*. (2013) 44:73–9. doi: 10.1161/STROKEAHA.112.670034

10. McKinley R, Häni L, Gralla J, El-Koussy M, Bauer S, Arnold M, et al. Fully automated stroke tissue estimation using random forest classifiers (FASTER). *J Cereb Blood Flow Metab*. (2017) 37:2728–41. doi: 10.1177/0271678X16674221

11. Nielsen A, Hansen MB, Tietze A, Mouridsen K. Prediction of tissue outcome and assessment of treatment effect in acute ischemic stroke using deep learning. *Stroke*. (2018) 49:1394–401. doi: 10.1161/STROKEAHA.117.019740

12. Robben D, Boers AMM, Marquering HA, Langezaal L, Roos Y, van Oostenbrugge RJ, et al. Prediction of final infarct volume from native CT perfusion and treatment parameters using deep learning. *Med Image Anal*. (2020) 59:101589. doi: 10.1016/j.media.2019.101589

13. Winzeck S, Hakim A, McKinley R, Pinto J, Alves V, Silva C, et al. ISLES 2016 and 2017-benchmarking ischemic stroke lesion outcome prediction based on multispectral MRI. *Front Neurol*. (2018) 9:679. doi: 10.3389/fneur.2018.00679

14. Benzakoun J, Charron S, Turc G, Hassen WB, Legrand L, Boulouis G, et al. Tissue outcome prediction in hyperacute ischemic stroke: comparison of machine learning models. *J Cereb Blood Flow Metab*. (2021) 41:3085–96. doi: 10.1177/0271678X211024371

15. Yu Y, Xie Y, Thamm T, Gong E, Ouyang J, Huang C, et al. Use of deep learning to predict final ischemic stroke lesions from initial magnetic resonance imaging. *JAMA Netw Open*. (2020) 3:e200772. doi: 10.1001/jamanetworkopen.2020.0772

16. Wang X, Fan Y, Zhang N, Li J, Duan Y, Yang B. Performance of machine learning for tissue outcome prediction in acute ischemic stroke: a systematic review and meta-analysis. *Front Neurol*. (2022) 13:910259. doi: 10.3389/fneur.2022.910259

17. Soltanpour M, Yousefnezhad M, Greiner R, Boulanger P, Buck B. Ischemic stroke lesion prediction using imbalanced temporal deep gaussian process (iTDGP). *arXiv*. (2022):221109068.

18. Kemmling A, Flottmann F, Forkert ND, Minnerup J, Heindel W, Thomalla G, et al. Multivariate dynamic prediction of ischemic infarction and tissue salvage as a function of time and degree of recanalization. *J Cereb Blood Flow Metab*. (2015) 35:1397–405. doi: 10.1038/jcbfm.2015.144

19. Winder AJ, Wilms M, Amador K, Flottmann F, Fiehler J, Forkert ND. Predicting the tissue outcome of acute ischemic stroke from acute 4D computed tomography perfusion imaging using temporal features and deep learning. *Front Neurosci*. (2022) 16:1009654. doi: 10.3389/fnins.2022.1009654

20. Amador K, Wilms M, Winder A, Fiehler J, Forkert ND. Predicting treatment-specific lesion outcomes in acute ischemic stroke from 4D CT perfusion imaging using spatio-temporal convolutional neural networks. *Med Image Anal*. (2022) 82:102610. doi: 10.1016/j.media.2022.102610

21. Forkert ND, Kaesemann P, Treszl A, Siemonsen S, Cheng B, Handels H, et al. Comparison of 10 TTP and Tmax estimation techniques for MR perfusion-diffusion mismatch quantification in acute stroke. *AJNR Am J Neuroradiol*. (2013) 34:1697–703. doi: 10.3174/ajnr.A3460

22. Bendszus M, Fiehler J, Subtil F, Bonekamp S, Aamodt AH, Fuentes B, et al. Endovascular thrombectomy for acute ischaemic stroke with established large infarct: multicentre, open-label, randomised trial. *Lancet*. (2023) 402:1753–63. doi: 10.1016/S0140-6736(23)02032-9

23. Ronneberger O, Fischer P, Brox T. *U-net: Convolutional networks for biomedical image segmentation*. Cham: Springer International Publishing (2015).

24. Alegiani AC, Dorn F, Herzberg M, Wollenweber FA, Kellert L, Siebert E, et al. Systematic evaluation of stroke thrombectomy in clinical practice: the German stroke registry endovascular treatment. *Int J Stroke*. (2019) 14:372–80. doi: 10.1177/1747493018806199

25. Wollenweber FA, Tiedt S, Alegiani A, Alber B, Bangard C, Berrouschot J, et al. Functional outcome following stroke thrombectomy in clinical practice. *Stroke*. (2019) 50:2500–6. doi: 10.1161/STROKEAHA.119.026005

26. Zaidat OO, Yoo AJ, Khatri P, Tomsick TA, Rv K, Saver JL, et al. Recommendations on angiographic revascularization grading standards for acute ischemic stroke. *Stroke*. (2013) 44:2650–63. doi: 10.1161/STROKEAHA.113.001972

27. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*. (2006) 31:1116–28. doi: 10.1016/j.neuroimage.2006.01.015

28. Rajashekar D, Wilms M, MacDonald ME, Ehrhardt J, Mouches P, Frayne R, et al. High-resolution T2-FLAIR and non-contrast CT brain atlas of the elderly. *Sci Data*. (2020) 7:56. doi: 10.1038/s41597-020-0379-9

29. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *NeuroImage*. (2012) 62:782–90. doi: 10.1016/j.neuroimage.2011.09.015

30. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, et al. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*. (2009) 45:S173–86. doi: 10.1016/j.neuroimage.2008.10.055

31. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*. (2008) 12:26–41. doi: 10.1016/j.media.2007.06.004

32. He K, Zhang X, Ren S, Sun J, editors. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; (2016). 27–30.

33. Hu J, Shen L, Sun G, editors. Squeeze-and-Excitation Networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*; (2018). 18–23 June 2018.

34. Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. *Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations*. Cham: Springer International Publishing (2017).

35. Misra D, editor Mish: A Self Regularized Non-Monotonic Activation Function. *British Machine Vision Conference*; (2020).

36. Kingma DP, Ba J. Adam: a method for stochastic optimization. *CoRR*. (2014):abs/1412.6980.

37. Stigler SM. Regression towards the mean, historically considered. *Stat Methods Med Res*. (1997) 6:103–14. doi: 10.1177/096228029700600202

38. Sharma M, Fox AJ, Symons S, Jairath A, Aviv RI. CT angiographic source images: flow-or volume-weighted? *Am J Neuroradiol*. (2011) 32:359–64. doi: 10.3174/ajnr.A2282

39. Boers AMM, Jansen IGH, Beenen LFM, Devlin TG, San Roman L, Heo JH, et al. Association of follow-up infarct volume with functional outcome in acute ischemic stroke: a pooled analysis of seven randomized trials. *J Neurointerv Surg*. (2018) 10:1137–42. doi: 10.1136/neurintsurg-2017-013724

40. Boers AMM, Jansen IGH, Brown S, Lingsma HF, Beenen LFM, Devlin TG, et al. Mediation of the relationship between endovascular therapy and functional outcome by follow-up infarct volume in patients with acute ischemic stroke. *JAMA Neurol*. (2019) 76:194–202. doi: 10.1001/jamaneurol.2018.3661

41. Kniep H, Meyer L, Broocks G, Bechstein M, Austein F, McDonough RV, et al. How much of the outcome improvement after successful recanalization is explained by follow-up infarct volume reduction? *J Neurointerv Surg*. (2023):jnis-2023-020296. doi: 10.1136/jnis-2023-020296

Glossary

AE	Adverse event
AI	Artificial intelligence
aOR	Adjusted odds ratio
ASPECTS	Alberta Stroke Program Early CT Score
CI	Confidence interval
CNN	Convolutional neural network
Coeff	Coefficient
CTA	Computed Tomography Angiography
CTP	Computed Tomography Perfusion
CV	Cross validation
DSC	Dice similarity coefficient
ICH	Intracranial hemorrhage
LVO	Large vessel occlusion
mRS	Modified Rankin scale
MIP	Maximum intensity projection
MT	Mechanical thrombectomy
NIHSS	National Institutes of Health Stroke Scale
Q1	1st quartile
Q3	3rd quartile
rtPA	recombinant tissue plasminogen activator
SD	Standard deviation
SE Block	Squeeze and excite block
sICH	Symptomatic intracranial hemorrhage
mTICI	modified Thrombolysis in cerebral infarction scale

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 15 April 2024
DOI 10.3389/fneur.2024.1341252

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Elisa Gouvêa Bogossian,
Université libre de Bruxelles, Belgium
Zhongheng Zhang,
Sir Run Run Shaw Hospital, China

*CORRESPONDENCE
Qichuan Zhuge
✉ qc.zhuge@wmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 20 November 2023
ACCEPTED 28 February 2024
PUBLISHED 15 April 2024

CITATION
Li X, Zhang C, Wang J, Ye C, Zhu J and Zhuge Q (2024) Development and performance assessment of novel machine learning models for predicting postoperative pneumonia in aneurysmal subarachnoid hemorrhage patients: external validation in MIMIC-IV.
Front. Neurol. 15:1341252.
doi: 10.3389/fneur.2024.1341252

COPYRIGHT
© 2024 Li, Zhang, Wang, Ye, Zhu and Zhuge. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development and performance assessment of novel machine learning models for predicting postoperative pneumonia in aneurysmal subarachnoid hemorrhage patients: external validation in MIMIC-IV

Xinbo Li[1,2†], Chengwei Zhang[1,2†], Jiale Wang[1,2], Chengxing Ye[1,2], Jiaqian Zhu[2] and Qichuan Zhuge[1,2*]

[1]Department of Neurosurgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2]Wenzhou Medical University, Wenzhou, China

Background: Postoperative pneumonia (POP) is one of the primary complications after aneurysmal subarachnoid hemorrhage (aSAH) and is associated with postoperative mortality, extended hospital stay, and increased medical fee. Early identification of pneumonia and more aggressive treatment can improve patient outcomes. We aimed to develop a model to predict POP in aSAH patients using machine learning (ML) methods.

Methods: This internal cohort study included 706 patients with aSAH undergoing intracranial aneurysm embolization or aneurysm clipping. The cohort was randomly split into a train set (80%) and a testing set (20%). Perioperative information was collected from participants to establish 6 machine learning models for predicting POP after surgical treatment. The area under the receiver operating characteristic curve (AUC), precision-recall curve were used to assess the accuracy, discriminative power, and clinical validity of the predictions. The final model was validated using an external validation set of 97 samples from the Medical Information Mart for Intensive Care IV (MIMIC-IV) database.

Results: In this study, 15.01% of patients in the training set and 12.06% in the testing set with POP after underwent surgery. Multivariate logistic regression analysis showed that mechanical ventilation time (MVT), Glasgow Coma Scale (GCS), Smoking history, albumin level, neutrophil-to-albumin Ratio (NAR), c-reactive protein (CRP)-to-albumin ratio (CAR) were independent predictors of POP. The logistic regression (LR) model presented significantly better predictive performance (AUC: 0.91) than other models and also performed well in the external validation set (AUC: 0.89).

Conclusion: A machine learning model for predicting POP in aSAH patients was successfully developed using a machine learning algorithm based on six perioperative variables, which could guide high-risk POP patients to take appropriate preventive measures.

KEYWORDS

aneurysmal subarachnoid hemorrhage, postoperative pneumonia, machine learning, endovascular treatment, prediction

Introduction

Aneurysmal subarachnoid hemorrhage (aSAH), primarily caused by the rupture of intracranial aneurysms, is a devastating disease with high morbidity and mortality (1, 2). The total occurrence of aSAH is about 9.1 per 100,000 people, accounts for 2–7% of all strokes (3, 4). Despite the advances in the treatment of aneurysmal subarachnoid hemorrhage (aSAH), such as endovascular coil embolization and microsurgery, postoperative complications can significantly impact the prognosis of patients, prolonged the time of hospitalization and lead to increased economic costs (2, 5).

Among these complications, postoperative pneumonia (POP), which occurs in 10–30% aSAH patients after surgical treatment, is considered a critical complication closely associated with the prognosis of aSAH patients (2, 6–10). Our previous study have found a close relationship between smoking history, delayed cerebral ischemia (DCI), mechanical ventilation time (MVT), Glasgow Coma Scale (GCS), Albumin, c-reactive protein (CRP). Using these variables, our team initially constructed a nomogram model to predicat POP for aSAH patients (11). But with the further research, numerous scholars contend many composite indicators were very effective for postoperative complications prediction, such as neutrophil-to-lymphocyte Ratio (NLR), prognostic nutrition index (PNI), neutrophil-to-albumin Ratio (NAR), CRP-to-albumin ratio (CAR), D-Dimer-to-Albumin Ratio (DAR) (2, 5, 12, 13). The aforementioned indicators, however, were not incorporated as variables in our previous study. The evaluation of the validity of these variables was therefore deemed necessary in order to ascertain their potential for enhancing the predictive power of our model.

In addition, although previous studies have suggested that conventional LR can provide a clinical prediction model that is easy to interpret, when conventional LR is used for complex multivariate non-linear relationships, complex transformations are often required owing to low robustness and multicollinearity between variables (14). Therefore, recent work has highlighted the potential of machine learning (ML) algorithms for stroke-related complications in stroke patients (15, 16). Savarraj et al. suggested that ML models significantly outperform conventional LR in predicting functional outcomes and has the potential to improve SAH management (17). The ML models possess the potential to outperform conventional linear or logistic regression models due to their exceptional ability in identifying intricate and nonlinear relationships among a multitude of prognostic variables (16). Thus, in this study, we conducted five ML prediction model, which contains support vector machine (SVM), logistic regression (LR), random forest (RF), multilayer perceptron (MLP), K-nearest neighbor (KNN) and extreme gradient boosting (XGBoost) for the prediction of POP within 30 days in aSAH patients after surgical treatment. The best model will be verified in the MIMIC-IV database. To supply clinical basis for the prevention and therapy of POP (18).

Materials and methods

Study population

Data for aged more than 18 years aSAH patients treated with intracranial aneurysm embolization or aneurysm clipping in the First Affiliated Hospital of Wenzhou Medical University from 1 June 2017 to 4 February 2022 were retrospectively collected. Patients will be included in this study if they meet the following requirements: (1) aged 18 years or older, who suffered their first SAH ever, admitted to our hospital within 24 h of symptom onset; (2) All aSAH patients should be confirmed by computed tomographic (CT), computed tomographic angiography (CTA) and digital subtraction angiography (DSA); (3) endovascular coiling of the aneurysm was performed; and (4) The definition of POP in this study refers to lower respiratory tract infections that occur within 30 days after endovascular coiling procedure in our hospital and the diagnosis of POP should follow modified Centers for Disease Control and Prevention (CDC) criteria, which was characterized by the following criteria: (1) A probable case of POP cannot be diagnosed based solely on the admission or follow-up chest x-ray, and it cannot be attributed to another diagnosis. (2) A proven case of POP is confirmed when there is a documented change in diagnosis observed on at least one chest x-ray image. For the purpose of this study, patients meeting the modified CDC criteria for probable/proven pneumonia were considered as cases, while those with pre-existing pneumonia prior to admission were excluded from analysis (19). A total of 893 patients met the above conditions, among which 187 patients were excluded for the following reasons: (1) 115 patients with other potential causes of SAH (96 patients with arteriovenous malformation, 10 patients with craniocerebral trauma, and 9 patients with hypertensive intracerebral hemorrhage); (2) 14 patient with history of malignant tumors, severe heart, hepatic or renal failure; (3) 7 patient previous use of antibiotics, systemic glucocorticoids, immunosuppressive agents, or immunotherapy within 1 month before admission; (4) proven POP had a confirmed change in diagnosis on at least one image of the chest x-ray, and 35 patients with pneumonia before admission were excluded (2, 19); (5) 14 aSAH patients who died within 24 h after surgery or lack complete case records were excluded. Finally, 706 patients were included in the internal cohort.

The external cohort was extracted from the Medical Information Mart for Intensive Care IV (MIMIC-IV) (version 2.2) database, derived from a large, freely accessible critical care database comprising 299,712 patients who were admitted to the ICU or the emergency department of Beth Israel Deaconess Medical Center between 2008 and 2019[1] (11). For the external cohort, individuals who met the following criteria were included: (1) those aged 18 years or older and

1 https://mimic.physionet.org/

(2) those with a diagnosis matching the International Classification of Diseases (ICD) code associated with aSAH among MIMIC-IV (ICD-9 code 430, and ICD-10 codes I60 to I609). A total of 1,172 patients fulfilled these conditions. Exclusion criteria consisted of: (1) non-first admission cases ($n = 33$); (2) history of malignant tumors, severe heart, hepatic or renal failure ($n = 120$); (3) length of stay less than 24 h in hospitalization ($n = 84$); and (4) patients without clinical data available for analysis ($n = 838$). Ultimately, a validation set comprising of 97 patients was selected from this group to validate the ML model. Figure 1 provides an overview of this process.

Variable selection

Collected variables included: (1) patients demographics (age, gender, history of smoking, alcoholics); (2) GCS, Hunt-Hess grades, modified Fisher (mFS) grade and WFNS grade on admission; (3) past medical history (hypertension, diabetes mellitus), coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), Stent assisted endovascular treatment (EVT), with craniotomy; (4) aneurysm location [anterior communicating artery (ACoA), internal carotid artery (ICA), middle cerebral artery (MCA), posterior communicating artery (PCoA), vertebrobasilar artery (VBA)]; (5) laboratory results were obtained within 24 h after admission in the context of a first examination [albumin, hemoglobin, neutrophils, monocytes, lymphocytes, uric acid, total cholesterol, triglycerides, neutrophils, monocytes, mean corpuscular volumec-reactive (MCV), CRP], and the composite index calculated by these laboratory results NLR = Neutrophil counts (*10^9 /L)/Lymphocyte count (*10^9 /L), PNI = Albumin (g/L) +10*Lymphocyte count (*10^9 /L), NAR = Neutrophil counts (*10^9 /L)/Albumin (g/L), CAR = CRP (mg/L) / Albumin (g/L), D-Dimer/Albumin Ratio (DAR) = D-dimer level (μg/mL)/Albumin (g/L).

The developed ML model of risk factors associated with POP in aSAH patients is relied on statistically significant results obtained through LASSO regression and multivariate logistic regression analysis ($p < 0.05$) (20).

ML model construction

A total of 706 patients with aSAH from our center were enrolled. We randomly divided the patients as training cohort ($N = 565$) and validation cohort ($N = 141$) according to a ratio of 80–20%. The training cohort was utilized to develop Linear models (logistic regression (LR), support vector machine (SVM)) and

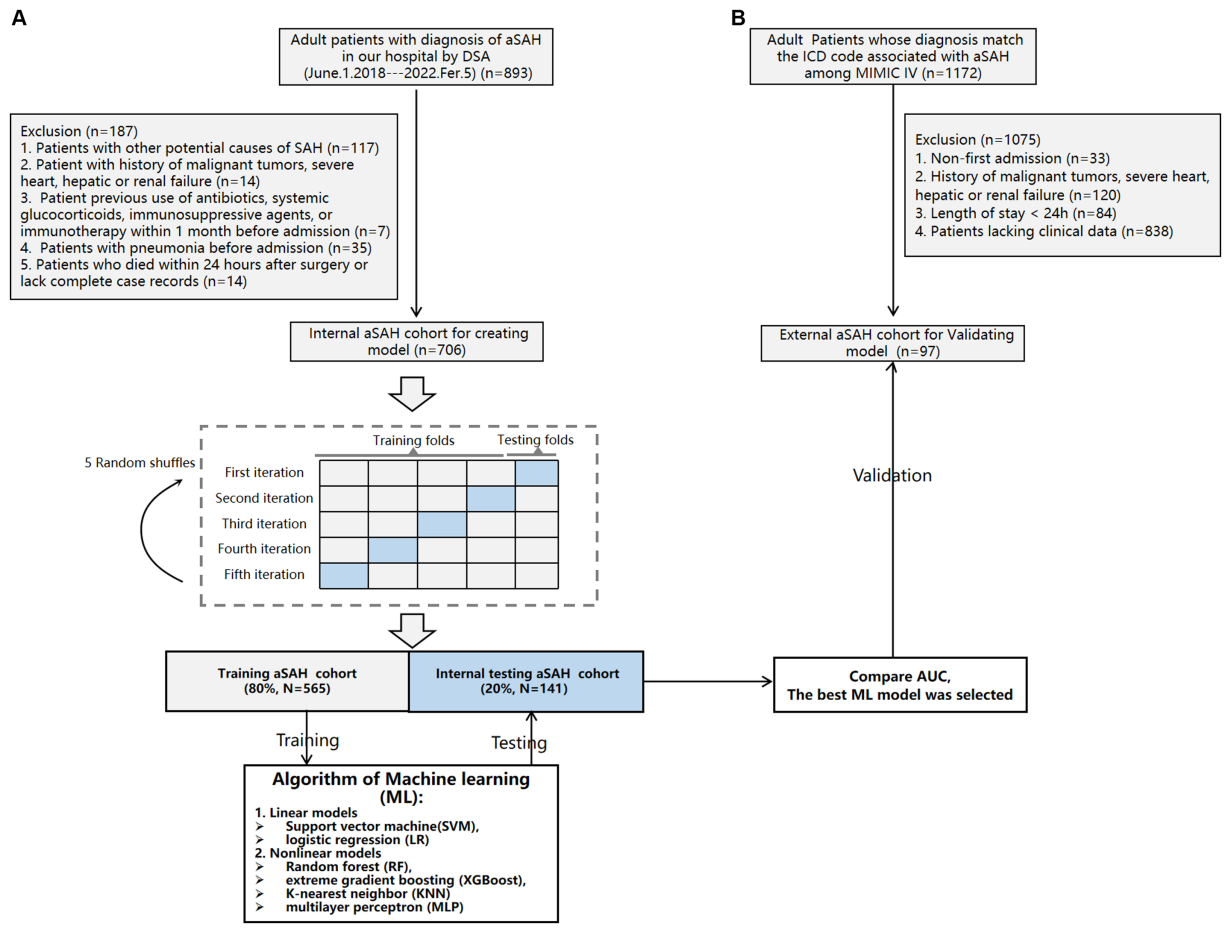


FIGURE 1
Flowchart of the internal cohort **(A)** and external cohort **(B)**.

Nonlinear models [XGBoost, k-nearest neighbor (KNN), random forest (RF), and multilayer perceptron (MLP)] (21–23). XGBoost model was constructed using the xgboost package.[2] The remaining five models were established via Scikit-learn package.[3] To develop an unbiased assessment of model performance, we performed 5 random shuffles of 5-fold cross-validation, as shown in Figure 1. Each iteration used a different stratified fold for model evaluation, and the remaining folds were used for model training (24). Subsequently, we recorded Area Under The Curve (AUC) to compare each ML models. Data processing and the ML process are summarized in Figure 1.

After the model was established, the SHapley Additive exPlanations (SHAP) package in Python was used to explain the model by analyzing two cases. The SHAP package interpreted the output of the machine learning model using a game-theoretic approach (25). SHA*p* values quantify the association of a variable with the outcome of a single patient, and the mean absolute SHAP value across all patients is reported as the SHAP value of avariable (26).

Statistical analysis

Continuous variables are represented in terms of mean and standard deviation (SD), while categorical variables are expressed in terms of frequency and percentage. Normally distributed and non-normally distributed variables were presented in the form of mean ± standard deviation and median (interquartile range). LASSO regression model was used to deal with collinearity of candidate variables, and the optimal predictive variable was selected (27). A multivariate logistic regression analysis was generated using predictors selected from the lasso analysis. Te features were represented by odds ratio (OR) and 95% confidence interval (CI). A two-tailed P value <0.05 was considered statistically significant. The stability of the columns in the validation queue is calculated using 1,000 boot replicas and a relative corrected C-index is calculated. All Statistical analyses were performed using R version 3.6.3 and python version 3.7.

Results

Demographic characteristics

Table 1 shows the clinical characteristics of the study population. A total of 706 aSAH patients were included in this study, which were divided into a training cohort ($N = 565$) and a validation cohort ($N = 141$). The number of patients with POP were 106 (16%) and 27 (12%) in training and testing cohorts, and men comprised 190 (34%) and 54 (38%) patients in the two groups, respectively. The median age in training cohort and testing cohorts were 55 and 59 years. The baseline data exhibited a high degree of consistency between the two groups ($p > 0.05$).

Feature selection

As partially relevant or less important features may negative affect performance of machine learning models, we performed feature selection by using LASSO regression. 39 variables with missing values <20% was extracted after interpolation, and 6 potential predictors were finally screened from the LASSO regression analysis (Figure 2). We included these 6 variables in multi-factor logistic regression and found that all 6 variables were statistically significant (Table 2). These 6 features were listed as follows: GCS (OR, 0.72; 95% CI, 0.58–0.73; $p < 0.001$), MVT (OR, 1.11; 95% CI, 1.03–1.21; $p = 0.01$), Albumin (OR, 0.90; 95% CI, 0.85–0.96; $p < 0.001$), NAR (OR, 55.23; 95% CI, 4.74–657.27; $p < 0.001$), CAR (OR, 2.61; 95% CI, 1.59–4.29; $p < 0.001$), Smoking history (OR, 8.37; 95% CI, 3.74–18.84; $p < 0.001$).

Machine learning model performance

Using the six features obtained by screening, we developed six machine learning models, including LR, SVM, RF, MLP, XGBoost, and KNN. Supplementary Table S1 and Figure 3 showed the best hyperparameter combination for each model and their AUCs in predicting POP. Their performance for prediction of POP was assessed (Table 3). The AUC values of KNN (0.78) and MLP (0.56) were relatively lower than LR (0.91), SVM (0.89), RF (0.87), XGBoost (0.86). Among them, LR exhibited the best performance for the prediction of POP risk. As the primary metric, the AUC for LR was 0.91 (95% confidence interval: 0.86–0.96). LR also exhibited the best performance based on the average precision of the precision-recall curve (0.65). The average precisions of the precision-recall curve for the remaining models are summarized in Figure 4. The remaining metrics are summarized in Table 3.

Application of the model

The SHAP package conducted a comprehensive analysis of training set, showing the impact of each variable on predicting POP (Figure 5). The preoperative and postoperative information of a patient was input into the model: with mechanical ventilation 5 days, the GCS score at admission was 8, albumin level 37.2 g/L, with smoke history, CAR 0.17, NAR 0.16. The model analyzed that the risk of POP in this patient was 85.0%, indicating that the probability of POP for the patients was high, and the implementation of preventive treatments against POP should be prioritized (Figure 6A). The preoperative and postoperative information of another patient was input into the model: with mechanical ventilation 2 days, the GCS score at admission was 14, albumin level 40.9 g/L, with no smoke history, CAR 0.10, NAR 0.12. The model analyzed that the risk of POP in this patient was 3.3%, which indicated a low probability of POP occurrence in the patients and the occurrence of POP in the patient does not require significant attention (Figure 6B). Furthermore, a website was established for clinicians to use the proposed LR model.[4]

2 https://xgboost.readthedocs.io/en/latest/python/index.html

3 https://github.com/scikit-learn/scikit-learn

4 https://www.xsmartanalysis.com/model/list/predict/model/html?mid=9730&symbol=4170wQvi00hz43760JQ1

TABLE 1 Characteristics of the study population.

Variable	Total (n = 706)	Training cohort (n = 565)	Testing cohort (n = 141)	p
Age, median [IQR]	56.00 [49.00, 66.00]	55.00 [48.00, 66.00]	59.00 [51.00, 66.00]	0.090
LOS, median [IQR]	14.00 [10.00, 19.00]	14.00 [10.00, 20.00]	13.00 [9.00, 19.00]	0.140
Gender (male), n (%)	244 (34.56)	190 (33.63)	54 (38.30)	0.297
Alcoholics, n (%)	90 (12.75)	78 (13.81)	12 (8.51)	0.092
Smoking history, n (%)	82 (11.61)	68 (12.04)	14 (9.93)	0.485
Hypertension, n (%)	353 (50.00)	288 (50.97)	65 (46.10)	0.300
CHD, n (%)	21 (2.97)	16 (2.83)	5 (3.55)	0.655
Diabetes, n (%)	51 (7.22)	41 (7.26)	10 (7.09)	0.946
COPD, n (%)	5 (0.71)	4 (0.71)	1 (0.71)	0.999
DCI, n (%)	41 (5.81)	36 (6.37)	5 (3.55)	0.199
Timing to DCI (day), mean (±SD)	6.1 ± 1.9	6.0 ± 2.0	6.3 ± 2.0	0.768
POP, n (%)	106 (15.01)	89 (15.75)	17 (12.06)	0.272
VBA aneurysm, n (%)	58 (8.22)	38 (6.73)	20 (14.18)	0.004
MCA aneurysm, n (%)	111 (15.72)	90 (15.93)	21 (14.89)	0.762
ICA aneurysm, n (%)	204 (28.90)	167 (29.56)	37 (26.24)	0.437
PCoA aneurysm, n (%)	149 (21.10)	119 (21.06)	30 (21.28)	0.955
ACoA aneurysm, n (%)	229 (32.44)	176 (31.15)	53 (37.59)	0.144
With Craniotomy, n (%)	120 (17.00)	94 (16.64)	26 (18.44)	0.610
Stent-assisted EVT, n (%)	303 (42.92)	245 (43.36)	58 (41.13)	0.633
GCS, median [IQR]	15.00 [14.00, 15.00]	15.00 [14.00, 15.00]	15.00 [15.00, 15.00]	0.603
Hunt-Hess grades, median [IQR]	2.00 [2.00, 2.00]	2.00 [2.00, 2.00]	2.00 [2.00, 2.00]	0.523
WFNS, median [IQR]	1.00 [1.00, 2.00]	1.00 [1.00, 2.00]	1.00 [1.00, 1.00]	0.632
mFS, median [IQR]	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	1.00 [1.00, 2.00]	0.103
MVT, day, median [IQR]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.104
Albumin (g/L), median [IQR]	38.90 [36.00, 41.50]	39.00 [36.00, 41.50]	38.70 [36.00, 41.40]	0.628
Glucose (mmlo/L), median [IQR]	6.40 [5.40, 7.60]	6.40 [5.40, 7.60]	6.60 [5.20, 7.80]	0.665
Triglyceride (μmlo/L), median [IQR]	1.07 [0.82, 1.64]	1.08 [0.82, 1.69]	1.05 [0.81, 1.50]	0.388
Uric acid (μmlo/L), median [IQR]	242.00 [176.00, 301.00]	241.00 [176.00, 301.00]	249.00 [174.00, 302.00]	0.414
Total cholesterol (μmlo/L), median [IQR]	4.87 [4.30, 5.64]	4.90 [4.30, 5.66]	4.83 [4.02, 5.59]	0.197
Neutrophil counts (*10^9/L), median [IQR]	9.82 [6.97, 12.70]	9.60 [6.96, 12.54]	10.05 [7.00, 12.70]	0.334
Monocyte count (*10^9/L), median [IQR]	0.47 [0.30, 0.70]	0.47 [0.31, 0.71]	0.47 [0.27, 0.70]	0.319
Lymphocyte count (*10^9 /L), median [IQR]	1.13 [0.86, 1.54]	1.14 [0.88, 1.55]	1.08 [0.72, 1.49]	0.258
Hemoglobin (g/L), median [IQR]	132.00 [121.00, 142.00]	132.00 [120.00, 142.00]	131.00 [121.00, 141.00]	0.574
MCV, (fl), median [IQR]	89.50 [86.40, 92.60]	90.00 [86.80, 92.70]	88.70 [86.00, 91.00]	0.003
Blood platelet count (*10^9/L), median [IQR]	213.00 [175.00, 257.00]	214.00 [175.00, 258.00]	211.00 [174.00, 254.00]	0.584
CRP (mg/L), median [IQR]	6.00 [2.60, 15.70]	6.00 [2.90, 15.70]	5.60 [1.30, 17.40]	0.298
D-dimer level (μg/mL), median [IQR]	1.14 [0.55, 2.48]	1.15 [0.54, 2.56]	1.10 [0.56, 2.19]	0.196
PNI, median [IQR]	45.05 [41.80, 48.30]	45.10 [41.80, 48.50]	44.90 [41.80, 47.60]	0.518
NAR, median [IQR]	0.25 [0.19, 0.33]	0.25 [0.19, 0.32]	0.25 [0.20, 0.34]	0.279
CAR, median [IQR]	0.16 [0.07, 0.42]	0.17 [0.07, 0.41]	0.14 [0.04, 0.45]	0.296
NLR, median [IQR]	8.55 [5.16, 13.75]	8.48 [5.11, 13.75]	8.74 [6.22, 13.75]	0.337
DAR, median [IQR]	2.96 [1.42, 6.49]	3.00 [1.38, 6.70]	2.74 [1.47, 6.05]	0.209

LOS, length of stay; ACoA, anterior communicating artery; PCoA, posterior communicating artery; ICA, Internal Carotid Artery; MCA, middle cerebral artery; VBA, vertebrobasilar aneurysm; POP, postoperative pneumonia; CRP, C-reactive protein; MCV, Mean Corpuscular Volume; GCS, Glasgow Coma Scale; WFNS, World Federation of Neurosurgical Societies; Stent-assisted EVT, Stent assisted endovascular treatment; DCI, delayed cerebral ischemia; PNI, prognostic nutrition index; NAR, neutrophil-to-albumin Ratio; NLR, neutrophil-to-lymphocyte ratio; CAR, C-reactive protein-to-albumin ratio; DAR, D-Dimer-to-Albumin Ratio; MVT, mechanical ventilation time.

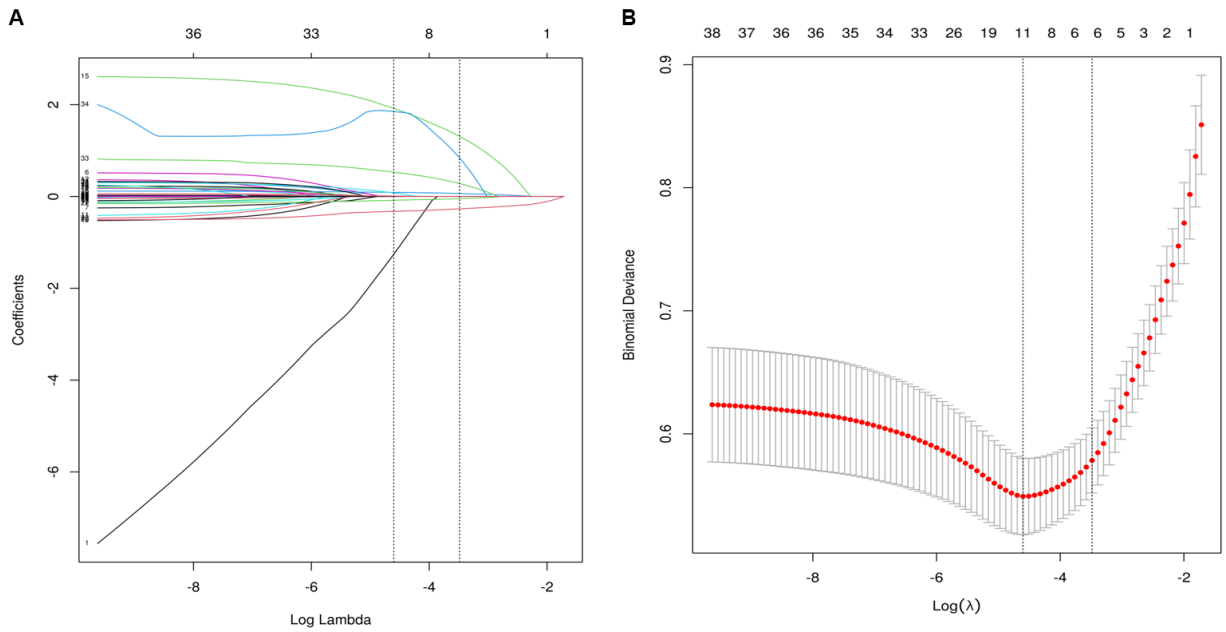


FIGURE 2
Perioperative variable selection using a LASSO logistic regression model. **(A)** The minimum criteria (lambda.min) and 1 SE of the minimum criteria (lambda. 1se) were used to depict the optimal values with dotted vertical lines. **(B)** LASSO coefficient profile of 39 variables. The coefficient profile is plotted according to the logarithmic sequence. To determine the optimal predictors of the model, five-fold cross-validation with minimum criteria was used, resulting in seven features with nonzero coefficients.

TABLE 2 **Multivariable Logistic Regression Model for Predicting postoperative pneumonia in aSAH patients.**

Predictor	β	SE	*p* value	Odds ratio (95% CI)
Intercept	5.14	2.13	0.460	196.4 (10.72–4054.43)
GCS, per score	−0.42	0.06	<0.001	0.72 (0.58–0.73)
MVT, per day	0.11	0.04	0.010	1.11 (1.03–1.21)
Albumin, g/L	−0.10	0.03	<0.001	0.90 (0.85–0.96)
NAR, per score	4.01	1.25	<0.001	55.23 (4.74–657.27)
CAR, per score	0.96	0.25	<0.001	2.61 (1.59–4.29)
Smoking history (yes vs. no)	2.12	0.41	<0.001	8.37 (3.74–18.84)

MVT, mechanical ventilation time; GCS, Glasgow Coma Scale; NAR, neutrophil-to-albumin Ratio; CAR, CRP-to-albumin ratio.

External data validation

To further confirmed the applicability of our model, we performed external validation using data from 97 patients with aSAH with MIME-IV. Figure 7 showed that the AUC of the model in the external data is 0.89 (95% confidence interval: 0.80–0.98), indicating that the model can still play a very good predictive performance in the external data.

Discussion

Early detection of POP is critical for timely interventions to prevent the onset of the complication (21). Numerous studies have developed ML models to predict postoperative pulmonary complications. Jong Ho Kim et al. developed an ML model to predict POP in patients undergoing surgery (28). Peng et al. had successfully created and verified a deep-neural-network model based on combined natural language data and structured data to predict pulmonary complications in geriatric patients (29). However, there is no POP prediction ML model designed explicitly for aSAH patients. Therefore, it is urgent to establish a POP clinical prediction model for aSAH patients, which has good application value in clinical identification and decision-making.

We used ML to develop models for the prediction of POP for aSAH patients. Model training using data from 565 patients was followed by model testing using data from 141 patients. Six algorithms (LR, XGBoost, RF, MLP, SVM, KNN) were used to develop the models, whereas four metrics were used to evaluate their performances. LR exhibited the best overall performance, with a specificity of 78% and a sensitivity of 94% in predicting POP in aSAH patients. Besides, the AUC values of MLP and KNN were relatively lower than XGBoost, RF and SVM, whose accuracy and robustness might be attributed to their nature of integrating multiple base

classifiers or learners (21). In addition, two examples were used to visualize how the LR model could predict POP and determine the relative importance of each variable for the clinician. With millions of POP taking place each year, the findings could help surgeons perform the management of postoperative pulmonary infection, such as the consideration of ventilator use time and tracheotomy, the use of prophylactic antibiotics. Finally, In order to reflect the model has a extensive range of applications, LR model was also validated on external data (MIMIC-IV) and showed good predictive performance.

In addition to successfully constructing machine learning models for predicting patient outcomes, the selection of variables should also prioritize the clinical requirements. As widely acknowledge, feature selection is a crucial process in machine learning. Selecting the proper combination of features to achieve a balance between model performance and efficiency is difficult but of great significance (30). Previous ML models required many features to predict Prognosis of patients, which reduces its practicality (28). We identified this issue at an early stage, prompting us to initially conduct LASSO regression on 39 variables. This approach not only mitigated the problem of collinearity among variables, but also resulted in a reduction of included variables in the model to six. Therefore, our ML models incorporated 6 variables of GCS, smoking history, MVT, GCS, NAR, CAR as predictors. These variables can be easily collected in clinical practice and important implications in clinical practice.

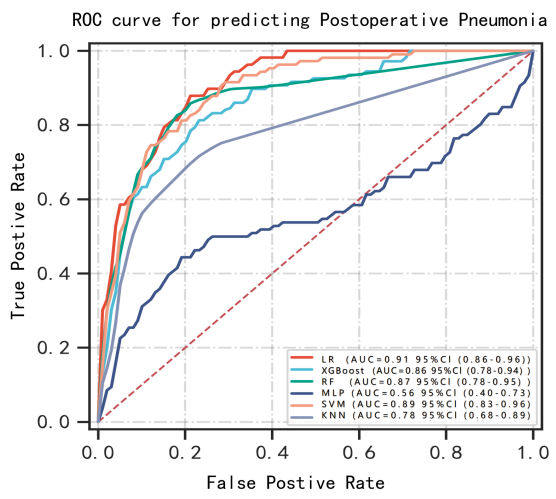


FIGURE 3
ROC curves for prediction of postoperative pneumonia (POP) on the test data set. Greater AUC shows higher discriminative ability of the model.

Among these variables, CAR was identified for the first time as an independent risk factor for predicting the POP in patients with aSAH. Previously, Dingding Zhang et al. discovered that an elevated CAR was correlated with the WFNS grade and Glasgow Outcome Scale (GOS) after 3 months aSAH (31). Our study further confirmed the importance of CAR on the prognosis of aSAH, and we also found that CAR was a better predictor of POP than other composite indicators such as NLR, NAR, DAR, and PNI (Supplementary Figure S1). Although Xin Zhang and Manman Xu et al. had demonstrated that NAR, PNI, DAR, and NLR are independent risk factors for predicting POP separately (2, 5, 12, 13), if these factors are compared together, the AUC of CAR and NAR was better than that of NLR,DAR and PNI, which also makes CAR and NAR remain as predictors into our model, while other indicators were excluded due to lower AUC or too high correlation coefficient with CAR and NAR. Back to the nature of the CAR, many previous studies had demonstrated the role of CRP and albumin in predicting pneumonia. CRP, an acute-phase protein, is triggered by different cytokines in reaction to infection, ischemia, trauma, and other inflammatory circumstances (32). In a recent study, Ben Gaastra et al. also discovered that CRP serves as an independent prognostic indicator for outcome following aSAH. The incorporation of CRP into prognostic models enhances their predictive accuracy (33). Xinlong Ma et al. found early increase in blood CRP appears to correlate with poor functional outcome after aSAH (34). Such patients exhibit a protracted recovery period for cough, expectoration, and swallowing function, necessitating an extended bed rest duration and presenting an increased susceptibility to pulmonary infections. Among them, the level of albumin played a pivotal role in our model, as it not only functioned as an independent prognostic indicator but also contributed to the calculation of two crucial predictive indicators, CAR and NAR. We posit that the primary factor contributing to this phenomenon is the multifaceted impact of hypoalbuminemia on patients with aSAH. Hypoalbuminemia in patients with aSAH not only reduces the level of immune protein and obstructs the repair of the mucosal barrier, leading to increased susceptibility to infection, but also exacerbates brain edema after subarachnoid hemorrhage, ultimately worsening motor function impairment in patients (32, 35). Some recent experimental findings additionally indicate that albuminemia mediates its neuroprotection through neurovascular remodeling and reducing cerebral lesions (36–38). A pilot study revealed that 1.25 g/kg/day albumin treatment was safe

TABLE 3 Performance of the six ML models in the testing set.

ML models	AUC (95%CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
LR	0.91 (0.86–0.96)	0.81 (0.78–0.84)	0.94 (0.91–0.98)	0.78 (0.69–0.87)
XGBoost	0.86 (0.78–0.94)	0.87 (0.85–0.88)	0.81 (0.67–0.96)	0.81 (0.73–0.89)
RF	0.87 (0.78–0.95)	0.87 (0.84–0.90)	0.83 (0.76–0.90)	0.85 (0.79–0.90)
MLP	0.56 (0.40–0.73)	0.76 (0.74–0.78)	0.48 (0.39–0.57)	0.80 (0.76–0.85)
SVM	0.89 (0.83–0.96)	0.79 (0.77–0.81)	0.87 (0.85–0.89)	0.81 (0.74–0.88)
KNN	0.78 (0.68–0.89)	0.85 (0.83–0.88)	0.70 (0.57–0.83)	0.81 (0.72–0.91)

LR, logistic regression; XGBoost, extreme gradient boosting; RF, random forest; MLP, multilayer perceptron; SVM, support vector machine; KNN, K-nearest neighbor.

in SAH patients and might produce a better outcome (37). The clinicians should therefore prioritize the monitoring of hypoalbuminemia in aSAH patients, and consider implementing albumin supplementation as a preventive measure against infection. There are few clinical studies in this area, and further research is warranted to establish more advanced clinical diagnostic and treatment protocols.

In addition, we also note that Xiao Jin et al. constructed a nomogram model for predicting POP in aSAH patients (11). All of our models use MIMIC-IV data as our external data to verify model efficacy. Xiao Jin's nomogram model had a validation performance of AUC 0.85 on external data, slightly lower than our LR model AUC 0.89. This is further evidence that, ML can offer unique perspective on the patient's condition and can serve as a decision support tool in the management of aSAH. However, clinical judgment is necessary to interpret the ML results and implement a corresponding plan of action (17). For example, in the process of modeling, because our database only had 5 patients with COPD, and these patients had mild intracerebral hemorrhage, and none of them had POP, COPD was not included as an important predictor in the process of modeling. However, according to clinical experience, COPD patients are extremely prone to pulmonary infection. Therefore, the knowledge and experience of doctors are essential for the prediction of POP in aSAH patients with special conditions.

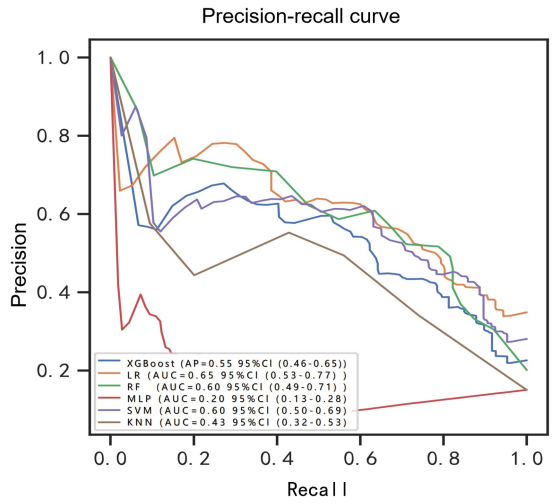


FIGURE 4
Precision-recall curve for prediction of postoperative pulmonary risk. BRF, balanced random forest; CI, confidence interval; LR, logistic regression; XGBoost, extreme gradient boosting; RF, random forest; MLP, multilayer perceptron; SVM, support vector machine; KNN, K-nearest neighbor.

This study had several limitations. First, although the ML models have been validated in another database, the ML models are developed on the basis of a single-center cohort study, and future multi-center study will be needed for external validation. Furthermore, it should be noted that this study was conducted in a retrospective manner, which may have introduced collection and entry biases as well as residual confounding factors." Third, in the case of external cohort inclusion, determining whether patients have received surgical treatment is challenging due to insufficient data availability. Hence, surgery was not considered as a prerequisite for enrolling an external cohort (11). In addition, it poses a significant challenge to monitor the occurrence of POP due to premature discharge resulting from severe illness or financial constraints, making it arduous to assess the likelihood of these patients developing POP. This limitation is equally applicable when utilizing the MIMIC-IV database, which solely documents the presence or absence of POP during hospitalization. Nonetheless, we believe that any potential error stemming from this constraint remains within manageable boundaries, as demonstrated in Xiao Jin et al.'s study (11). Last but not least, two or more machine learning algorithms can be synthesized to further improve predictive accuracy. Such a process is referred to as ensemble modeling, and it has been used broadly in various industries (39). However, because the authors of

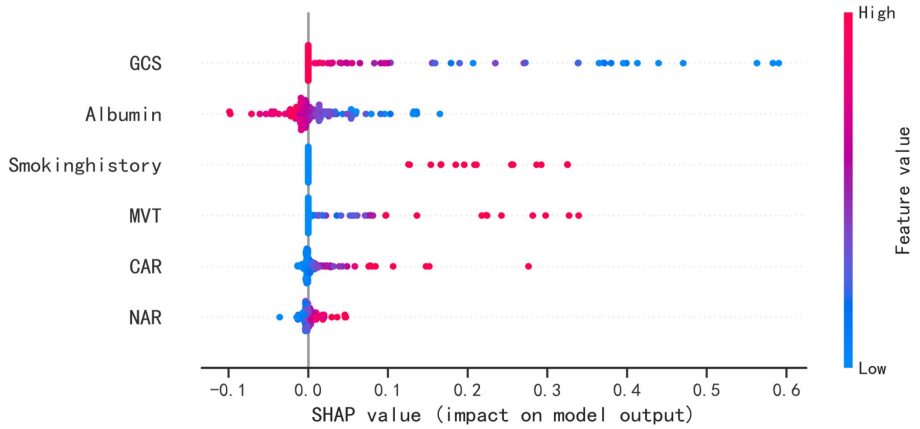


FIGURE 5
SHAP analysis of the proposed model on the testing set. This figure described data from the testing set, with each point representing one patient. The color represents the value of the variable; red represents the larger value; blue represents the smaller value. The horizontal coordinates represent a positive or negative correlation with transfusion risk, with a positive value indicating a risk of POP and a negative value indicating no risk for POP. The absolute value of the horizontal coordinate indicates the degree of influence; the greater the absolute value of the horizontal coordinate, the greater the degree of influence.

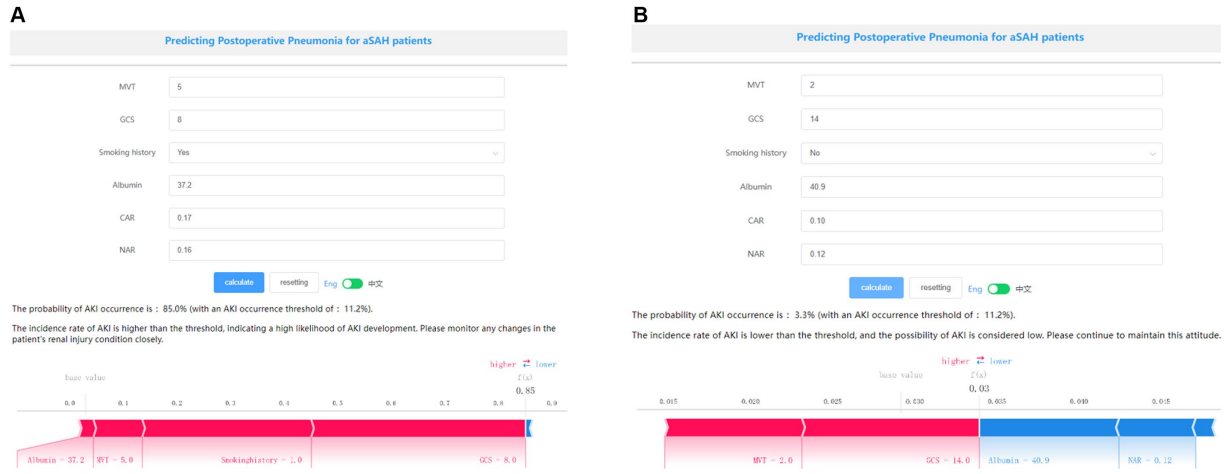


FIGURE 6
Examples of website usage. Entering the input value determined the transfusion requirements and displayed how each value contributed to the prediction. **(A)** Example 1 has a high probability of POP, the probability of POP is 85.0%, and **(B)** Example 2 has a low probability of POP, the probability of POP is 3.3%.

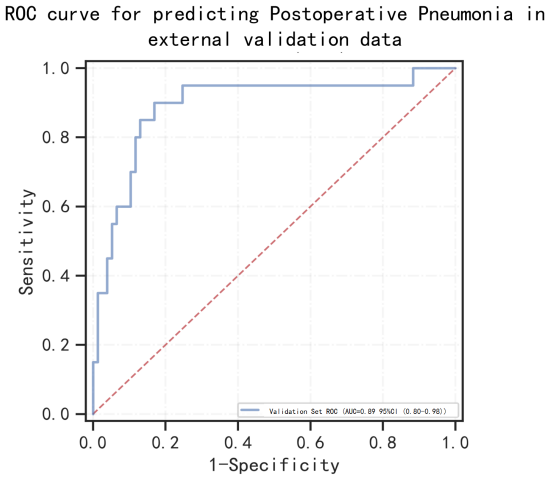


FIGURE 7
Receiver operating characteristic (ROC) curves for predicting postoperative pneumonia in aneurysmal subarachnoid hemorrhage patients in the external valication set.

this paper were unfamiliar with this method, they did not dare to operate ensemble modeling arbitrarily. In the future, we will pay more attention to improving the prediction performance of clinical models by ensemble learning algorithms.

Conclusion

Our study has successfully established six novel ML models to predict POP among aSAH patients. Of these, the LR model has demonstrated overall best performance. Furthermore, an online prediction tool based on the LR model was developed to identify patients at high risk for POP after aSAH and facilitate timely interventions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Committee in Clinical Research (ECCR) of the First Affiliated Hospital of Wenzhou Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

XL: Writing – original draft, Writing – review & editing. CZ: Writing – original draft, Writing – review & editing. JW: Formal analysis, Data curation, Methodology, Writing – original draft, Writing – review & editing. CY: Writing – original draft, Writing – review & editing. JZ: Data curation, Writing-orginal draft, Writing-review & editing. QZ: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study received funding from Wenzhou Science and Technology Bureau Major Project ZS2017007 and Key project of International Cooperation of National Natural Science Foundation 81820108011. This work was supported by Extreme Smart Analysis platform (https://www.xsmartanalysis.com/).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1341252/full#supplementary-material

References

1. Maher M, Schweizer TA, Macdonald RL. Treatment of spontaneous subarachnoid hemorrhage: guidelines and gaps. *Stroke*. (2020) 51:1326–32. doi: 10.1161/STROKEAHA.119.025997

2. Zhang X, Zhang S, Wang C, Liu R, Li A. High neutrophil-to-albumin ratio predicts postoperative pneumonia in aneurysmal subarachnoid hemorrhage. *Front Neurol*. (2022) 13:840858. doi: 10.3389/fneur.2022.840858

3. Aneurysmal D'SS, Hemorrhage S. Aneurysmal subarachnoid Hemorrhage. *J Neurosurg Anesthesiol*. (2015) 27:222–40. doi: 10.1097/ANA.0000000000000130

4. Ingall T, Asplund K, Mähönen M, Bonita R. A multinational comparison of subarachnoid hemorrhage epidemiology in the WHO MONICA stroke study. *Stroke*. (2000) 31:1054–61. doi: 10.1161/01.STR.31.5.1054

5. Xu M, Zhang L, Wang J, Cheng L, Chen C, Li S, et al. Pre-operative prognostic nutrition index and post-operative pneumonia in aneurysmal subarachnoid hemorrhage patients. *Front Neurol*. (2023) 14:1045929. doi: 10.3389/fneur.2023.1045929

6. Wu MR, Chen YT, Li ZX, Gu HQ, Yang KX, Xiong YY, et al. Dysphagia screening and pneumonia after subarachnoid hemorrhage: findings from the Chinese stroke center alliance. *CNS Neurosci Ther*. (2022) 28:913–21. doi: 10.1111/cns.13822

7. Hop JW, Rinkel GJ, Algra A, van Gijn J. Case-fatality rates and functional outcome after subarachnoid hemorrhage: a systematic review. *Stroke*. (1997) 28:660–4. doi: 10.1161/01.STR.28.3.660

8. Inagawa T. Trends in incidence and case fatality rates of aneurysmal subarachnoid hemorrhage in Izumo City, Japan, between 1980-1989 and 1990-1998. *Stroke*. (2001) 32:1499–507. doi: 10.1161/01.STR.32.7.1499

9. Ding CY, Peng L, Lin YX, Yu LH, Wang DL, Kang DZ. Elevated lactate dehydrogenase level predicts postoperative pneumonia in patients with aneurysmal subarachnoid hemorrhage. *World Neurosurg*. (2019) 129:e821–30. doi: 10.1016/j.wneu.2019.06.041

10. Li R, Lin F, Chen Y, Lu J, Han H, Yan D, et al. In-hospital complication-related risk factors for discharge and 90-day outcomes in patients with aneurysmal subarachnoid hemorrhage after surgical clipping and endovascular coiling: a propensity score-matched analysis. *J Neurosurg*. (2021) 137:381–92. doi: 10.3171/2021.10.JNS211484

11. Jin X, Wang S, Zhang C, Yang S, Lou L, Xu S, et al. Development and external validation of a nomogram for predicting postoperative pneumonia in aneurysmal subarachnoid hemorrhage. *Front Neurol*. (2023) 14:1251570. doi: 10.3389/fneur.2023.1251570

12. Cai L, Zeng H, Tan X, Wu X, Qian C, Chen G. The role of the blood neutrophil-to-lymphocyte ratio in aneurysmal subarachnoid hemorrhage. *Front Neurol*. (2021) 12:671098. doi: 10.3389/fneur.2021.671098

13. Wu W, Liu X, Zhu Q, Chen X, Sheng B, Zhang J, et al. The D-dimer/albumin ratio is a prognostic marker for aneurysmal subarachnoid Hemorrhage. *Brain Sci*. (2022) 12:1700. doi: 10.3390/brainsci12121700

14. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol*. (1996) 49:1225–31. doi: 10.1016/S0895-4356(96)00002-9

15. Alaka SA, Menon BK, Brobbey A, Williamson T, Goyal M, Demchuk AM, et al. Functional outcome prediction in ischemic stroke: a comparison of machine learning algorithms and regression models. *Front Neurol*. (2020) 11:889. doi: 10.3389/fneur.2020.00889

16. Mutke MA, Madai VI, Hilbert A, Zihni E, Potreck A, Weyland CS, et al. Comparing poor and favorable outcome prediction with machine learning after mechanical thrombectomy in acute ischemic stroke. *Front Neurol*. (2022) 13:737667. doi: 10.3389/fneur.2022.737667

17. Savarraj JPJ, Hergenroeder GW, Zhu L, Chang T, Park S, Megjhani M, et al. Machine learning to predict delayed cerebral ischemia and outcomes in subarachnoid hemorrhage. *Neurology*. (2021) 96:e553–62. doi: 10.1212/WNL.0000000000011211

18. Grimes DA. The nomogram epidemic: resurgence of a medical relic. *Ann Intern Med*. (2008) 149:273–5. doi: 10.7326/0003-4819-149-4-200808190-00010

19. Smith CJ, Kishore AK, Vail A, Chamorro A, Garau J, Hopkins SJ, et al. Diagnosis of stroke-associated pneumonia: recommendations from the pneumonia in stroke consensus group. *Stroke*. (2015) 46:2335–40. doi: 10.1161/STROKEAHA.115.009617

20. Wei X, Min Y, Yu J, Wang Q, Wang H, Li S, et al. The value of admission serological indicators for predicting 28-day mortality in intensive care patients with acute heart failure: construction and validation of a nomogram. *Front Cardiovasc Med*. (2021) 8:741351. doi: 10.3389/fcvm.2021.741351

21. Chen C, Yang D, Gao S, Zhang Y, Chen L, Wang B, et al. Development and performance assessment of novel machine learning models to predict pneumonia after liver transplantation. *Respir Res*. (2021) 22:94. doi: 10.1186/s12931-021-01690-3

22. Hu P, Li Y, Liu Y, Guo G, Gao X, Su Z, et al. Comparison of conventional logistic regression and machine learning methods for predicting delayed cerebral ischemia after aneurysmal subarachnoid Hemorrhage: a multicentric observational cohort study. *Front Aging Neurosci*. (2022) 14:857521. doi: 10.3389/fnagi.2022.857521

23. Lee HC, Yoon SB, Yang SM, Kim WH, Ryu HG, Jung CW, et al. Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. *J Clin Med*. (2018) 7:428. doi: 10.3390/jcm7110428

24. Xue B, Li D, Lu C, King CR, Wildes T, Avidan MS, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw Open*. (2021) 4:e212240. doi: 10.1001/jamanetworkopen.2021.2240

25. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. (2020) 2:56–67. doi: 10.1038/s42256-019-0138-9

26. van den Bosch T, Warps AK, de Nerée tot Babberich MPM, Stamm C, Geerts BF, Vermeulen L, et al. Predictors of 30-day mortality among Dutch patients undergoing colorectal cancer surgery, 2011-2016. *JAMA Netw Open*. (2021) 4:e217737. doi: 10.1001/jamanetworkopen.2021.7737

27. Liu J, Huang J, Ma S, Wang K. Incorporating group correlations in genome-wide association studies using smoothed group lasso. *Biostatistics*. (2013) 14:205–19. doi: 10.1093/biostatistics/kxs034

28. Kim JH, Kim Y, Yoo K, Kim M, Kang SS, Kwon YS, et al. Prediction of postoperative pulmonary edema risk using machine learning. *J Clin Med*. (2023) 12:1804. doi: 10.3390/jcm12051804

29. Peng X, Zhu T, Chen G, Wang Y, Hao X. A multicenter prospective study on postoperative pulmonary complications prediction in geriatric patients with deep neural network model. *Front Surg*. (2022) 9:976536. doi: 10.3389/fsurg.2022.976536

30. Shafiee M, Lotfi FH, Saleh H. Supply chain performance evaluation with data envelopment analysis and balanced scorecard approach. *Appl Math Model*. (2014) 38:5092–112. doi: 10.1016/j.apm.2014.03.023

31. Zhang D, Yan H, Wei Y, Liu X, Zhuang Z, Dai W, et al. C-reactive protein/albumin ratio correlates with disease severity and predicts outcome in patients with aneurysmal subarachnoid Hemorrhage. *Front Neurol*. (2019) 10:1186. doi: 10.3389/fneur.2019.01186

32. Bozada-Gutiérrez K, Trejo-Ávila M, Moreno-Portillo M. Postoperative complications and predictors of mortality in patients with COVID-19. *Cir Cir*. (2023) 91:344–53. doi: 10.24875/CIRU.22000512

33. Gaastra B, Barron P, Newitt L, Chhugani S, Turner C, Kirkpatrick P, et al. CRP (C-reactive protein) in outcome prediction after subarachnoid hemorrhage and the role of machine learning. *Stroke*. (2021) 52:3276–85. doi: 10.1161/STROKEAHA.120.030950

34. Ma X, Lan F, Zhang Y. Associations between C-reactive protein and white blood cell count, occurrence of delayed cerebral ischemia and poor outcome following

aneurysmal subarachnoid hemorrhage: a systematic review and meta-analysis. *Acta Neurol Belg*. (2021) 121:1311–24. doi: 10.1007/s13760-020-01496-y

35. Fakler JK, Grafe A, Dinger J, Josten C, Aust G. Perioperative risk factors in patients with a femoral neck fracture - influence of 25-hydroxyvitamin D and C-reactive protein on postoperative medical complications and 1-year mortality. *BMC Musculoskelet Disord*. (2016) 17:51. doi: 10.1186/s12891-016-0906-1

36. Xie Y, Liu W, Zhang X, Wang L, Xu L, Xiong Y, et al. Human albumin improves long-term behavioral sequelae after subarachnoid hemorrhage through neurovascular remodeling. *Crit Care Med*. (2015) 43:e440–9. doi: 10.1097/CCM.0000000000001193

37. Suarez JI, Martin RH, Calvillo E, Dillon C, Bershad EM, Macdonald RL, et al. The albumin in subarachnoid Hemorrhage (ALISAH) multicenter pilot clinical trial: safety and neurologic outcomes. *Stroke*. (2012) 43:683–90. doi: 10.1161/STROKEAHA.111.633958

38. Richter F, Mehdorn AS, Fedders T, Reichert B, Egberts JH, Becker T, et al. C-reactive protein as predictor for infectious complications after robotic and open esophagectomies. *J Clin Med*. (2022) 11:5654. doi: 10.3390/jcm11195654

39. Zhang Z, Chen L, Xu P, Hong Y. Predictive analytics with ensemble modeling in laparoscopic surgery: a technical note. *Laparosc Endosc Robot Surg*. (2022) 5:25–34. doi: 10.1016/j.lers.2021.12.003

frontiers | Frontiers in Neurology

TYPE Perspective
PUBLISHED 02 May 2024
DOI 10.3389/fneur.2024.1389056

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Veronica Olavarria,
Clínica Alemana, Chile
Andrew P. Reimer,
Case Western Reserve University,
United States

*CORRESPONDENCE
Stephen W. English
✉ english.stephen@mayo.edu

RECEIVED 20 February 2024
ACCEPTED 22 April 2024
PUBLISHED 02 May 2024

CITATION
Wolcott ZC and English SW (2024) Artificial intelligence to enhance prehospital stroke diagnosis and triage: a perspective. *Front. Neurol.* 15:1389056. doi: 10.3389/fneur.2024.1389056

COPYRIGHT
© 2024 Wolcott and English. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence to enhance prehospital stroke diagnosis and triage: a perspective

Zoe C. Wolcott and Stephen W. English*

Department of Neurology, Mayo Clinic, Jacksonville, FL, United States

As health systems organize to deliver the highest quality stroke care to their patients, there is increasing emphasis being placed on prehospital stroke recognition, accurate diagnosis, and efficient triage to improve outcomes after stroke. Emergency medical services (EMS) personnel currently rely heavily on dispatch accuracy, stroke screening tools, bypass protocols and prehospital notification to care for patients with suspected stroke, but novel tools including mobile stroke units and telemedicine-enabled ambulances are already changing the landscape of prehospital stroke care. Herein, the authors provide our perspective on the current state of prehospital stroke diagnosis and triage including several of these emerging trends. Then, we provide commentary to highlight potential artificial intelligence (AI) applications to improve stroke detection, improve accurate and timely dispatch, enhance EMS training and performance, and develop novel stroke diagnostic tools for prehospital use.

KEYWORDS

prehospital stroke, artificial intelligence, mobile stroke unit, telemedicine, computer vision, stroke triage, LVO stroke

Introduction

Despite significant advancements in stroke treatments including thrombolytic therapy and mechanical thrombectomy for patients with large vessel occlusion (LVO), stroke remains a leading cause of long-term disability (1). These therapies operate on the principle that recanalization of occluded arteries reduces brain injury, but treatment is highly time dependent (2). Moreover, access to stroke treatment remains limited, as few hospitals have resources to deliver these interventions. Therefore, healthcare systems have organized such that hospitals with primary stroke center (PSC) certification can administer thrombolytic treatment, whereas only comprehensive stroke centers (CSC) are accredited to perform mechanical thrombectomy in patients with LVO stroke (3, 4). This regional distribution increases pressure on our Emergency Medical Services (EMS) to ensure patients are directed to the appropriate hospital. The stakes are high, emphasized by a recent US stroke registry reported average door-in-door-out time of 174 min for patients required transfer to a higher level of care, and up to one-third of patients with LVO become ineligible for thrombectomy due to delays in transfer to a comprehensive stroke center (3–7). Therefore, EMS personnel have become crucial stakeholders in the stroke care continuum.

Current state

Prehospital EMS systems have adapted to the changing acute stroke treatment landscape. Emergency 9-1-1 dispatch protocols prioritize stroke calls as the highest acuity (8, 9). EMS

prehospital notification facilitates early mobilization of stroke teams and downstream resources (9, 10). And scales designed to identify LVO stroke have been validated to enhance prehospital triage (11, 12). National recommendations call for an initial stroke screening tool, and if positive, an LVO screening tool in patients within 24 h of last known well. Recommendations call for bypass of primary stroke centers if the comprehensive stroke center is within 30 min for urban, 45 min for suburban, and 60 min for rural settings if the prehospital LVO screen is positive (11). Despite these changes, many stroke patients are not recognized in the prehospital setting. Contributing factors include inadequate training, limited EMS resources, high EMS personnel turnover, suboptimal scale accuracy, poor scale utilization, and inherent variability in presenting stroke symptoms.

Despite these shortcomings, there are reasons for optimism. First, incorporating prehospital LVO screening tools have demonstrated almost 3-fold higher thrombectomy rates for eligible patients presenting directly to thrombectomy-capable centers (4.8% without vs. 13.6% with LVO screening tool) (3, 12). And recently, mobile stroke units (MSU)—specialized ambulances equipped with CT, telemedicine, and stroke nurses—have demonstrated both faster treatment (10-fold increase in patients treated within 60 min of symptom onset) and an 8% increase in excellent outcomes 90 days—compared to conventional ambulance care (13, 14).

To date, the MSU model has experienced limited uptake, likely driven by several factors: (1) high MSU cost (both upfront and recurring costs), (2) poor reimbursement, and (3) logistical and bureaucratic challenges to integrate and scale within existing EMS infrastructure. Due to these challenges, several groups have explored ambulance-based telemedicine to bring vascular neurology expertise to the ambulance (11, 15); one group found that ambulance-based telemedicine was 100% accurate in predicting reperfusion candidates compared to 70% for the LVO scale (11, 15). This approach has key advantages, as it is more scalable due to lower upfront costs, can leverage existing ambulances, and can be leveraged for non-neurologic indications including trauma and cardiac emergencies.

Future state: artificial intelligence to enhance prehospital stroke care

MSUs and telemedicine-enabled ambulances highlight how the paradigm shift for prehospital stroke care has promoted innovations to confront this emerging demand. We foresee the most impactful ways to minimize prehospital delays include improved public awareness, efficient and accurate emergency dispatch, enhanced paramedic evaluation and triage, and streamlined inter-facility workflows (16, 17). Future innovations in telemedicine and artificial intelligence (AI) will undoubtedly improve prehospital stroke diagnosis and triage and allow for more standardization in clinical practice. The authors provide insights into specific opportunities we believe AI applications will hold the most promise to enhance prehospital stroke care (Figure 1).

Stroke recognition tools

Perhaps the most common reason for delays in stroke treatment can be attributed to delayed recognition of stroke symptoms by patients or families. AI solutions via wearable technologies and computer vision may help reduce this barrier by expediting EMS activation (18). Wireless sensors and wearable technologies including watches, fitness bands, and rings could be used to characterize limb

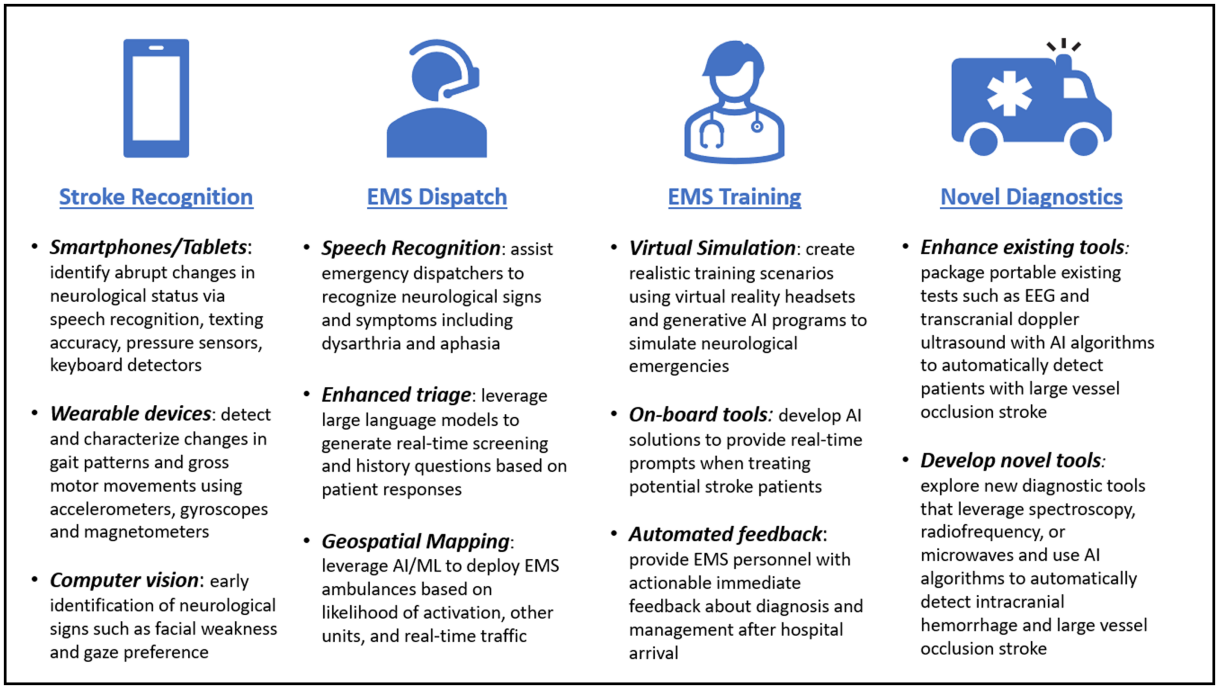


FIGURE 1
Overview of potential artificial intelligence applications to enhance prehospital stroke diagnosis and triage.

motor deficits and gait patterns using accelerometers, gyroscopes, and magnetometers (18, 19). Smartphones, tablets, and computers can be turned into powerful screening tools by identifying sudden changes in motor function, speech, or language by incorporating artificial intelligence in touch screen, keyboard, and speech-to-text inputs through pressure sensors, keyboard detectors, and speech recognition software (18). This technology has the potential to alert individuals to the onset of stroke symptoms and hasten EMS activation (19, 20). For example, one recent AI algorithm reportedly reviewed 269 patients and found 97% accuracy in detection of facial asymmetry and 72% accuracy in detection of arm weakness (21, 22), and improvements in eye tracking may identify sudden changes in eye movements or unilateral gaze deviation, a dependable biomarker for LVO stroke (23, 24). In fact, observation of head and/or gaze deviation alone in telemedicine consultations has been shown to predict LVO with a sensitivity of 0.70 and specificity of 0.93, and MT necessity with a sensitivity of 0.86 and specificity of 0.90 (23). Incorporating automated detection of gaze preference, aphasia, neglect, and facial palsy into a single application may aid in early detection of acute strokes and identify patients with LVO strokes, facilitating pre-hospital triage to a comprehensive stroke center (25).

EMS dispatch

Emergency dispatcher stroke recognition is associated with higher EMS utilization of stroke scales, EMS stroke identification, and faster response times (26–28). Despite specific stroke training, the overall accuracy of emergency dispatchers leaves room for improvement (26, 29). Several groups have investigated AI as a means to improve dispatch accuracy and resource allocation. One study reviewed the potential impact of automated speech recognition software on stroke recognition in Denmark, and found that it could improve stroke detection by 16% and increase thrombolysis use by 2% (30). Another study scrutinized the application of AI algorithms in ambulance dispatch protocols for motor vehicle accidents. Through predictive analytics, this study anticipated optimal deployment times and locations for ambulances to expedite response times (31). We envision a future state where similar AI strategies can be deployed to position ambulances based on factors such as time of day, regional demographics, and other conflicting emergency calls to ensure fast response.

EMS personnel training and competency assessments

Generative AI and virtual reality may soon transform medical education through the development of immersive medical simulation experiences (32, 33). Stroke only accounts for approximately 2% of total EMS dispatches, and with high turnover and limited shifts, EMS personnel may lack appropriate training or experience to identify and triage stroke. While screening tools are helpful, specific simulation exercises leveraging virtual or augmented reality can provide structured educational experiences to augment training prior to real-world encounters. We envision specific software tailored to each healthcare provider role that fosters practice on simulated stroke patients in a "no-stakes" environment. These tools can provide real-time feedback to enhance retention and give EMS personnel valuable repetitions to improve performance. Similar training experiences have already been developed for surgical training and sepsis management (34, 35). Moreover, real-time AI-enhanced feedback could significantly enhance EMS provider training and competence by offering immediate insights during patient assessment, helping them to learn and adapt more effectively (30, 35, 36). This ongoing training can lead to improved performance, ensuring our EMS personnel are equipped to handle the complexities of stroke diagnosis. Nonetheless, it's crucial to recognize that while AI can augment clinical scales and inform examinations, the inherent variability and frequency of stroke mimics leave room for error if triage decisions are solely reliant on clinical severity scores (34, 37).

Intelligent telemedicine

Stroke neurologists were pioneering end-users of telemedicine, connecting with remote patients to provide real-time audio-video consultation in remote underserved locations for over a decade (38). The COVID-19 pandemic has accelerated the uptake of telemedicine, creating opportunities to explore AI applications to the video encounter including computer vision tools that can precisely measure human movements from videos and may allow for autonomous recognition of stroke symptoms (39). Clear use cases for image-based AI applications have already been reported in the fields such as cardiology, pathology, dermatology, and ophthalmology (40), but by applying deep-learning algorithms to live patient videos, we believe this may soon allow for the automated assessment of neurologic examinations such as the LVO stroke scales and the NIHSS. While these applications have not been tested in this clinical context, we envision the use of future computer vision applications, either alone or in conjunction with video telestroke services to more accurately recognize focal neurologic deficits and improve prehospital triage.

Novel diagnostic tools

The ability to differentiate hemorrhagic stroke, non-LVO stroke, and LVO stroke remains a critical challenge in prehospital stroke care. AI integration with both existing and emerging technologies holds promise to develop cost-effective tools to assist in prehospital stroke diagnosis. Efforts in this arena are ongoing, including applications that harness electroencephalography (EEG) and transcranial doppler (TCD) to identify LVO. Both of these are attractive options due to their inherent portability, low-cost, and potential discriminatory ability to differentiate stroke syndromes. One study demonstrated an EEG cap during ambulance transportation can accurately differentiate LVO stroke, reporting a diagnostic accuracy of 91% and a positive likelihood ratio of 11.0 (41). Another study investigating a portable, automated TCD device and found the overall accuracy of LVO detection was 91% when combining two biomarkers. While these results demonstrated the potential value of AI to enhance existing point-of-care diagnostics, these devices have not been rigorously tested in a real-world prehospital setting yet and may not yet be ready for widespread use.

Advancements in AI have facilitated the development of novel diagnostic methods capable of swift detection and triage of stroke

patients. Automated headsets have demonstrated potential in characterizing brain tissue characteristics by utilizing spectroscopy, radiofrequency, and microwave detection (42–44). AI algorithms applied in this context have been successful in differentiating patients with acute intracerebral hemorrhage and LVO stroke from healthy controls (20, 43, 44). One emerging example is a Volumetric Impedance Phase Shift Spectroscopy (VIPS) headset, a non-invasive tool that measures electrical impedance of different tissue types using electromagnetic waves. VIPS can detect lateralized changes associated with brain injury, and research suggests it can accurately identify patients with severe stroke (either LVO stroke or large ICH) with a 93% sensitivity and 87% specificity (43). Similar headsets are investigating the role of radiofrequency and low-energy microwave signals couple with machine learning algorithms to detect traumatic brain injury, intracranial hemorrhage, and LVO stroke, but results have not yet been published (44). With time, we believe these devices may provide scalable solutions to enhance existing regional stroke triage.

Conclusion

As the landscape of acute stroke treatment continues to evolve, the significance of prehospital stroke recognition, accurate diagnosis, and effective triage has become increasingly valuable to ensuring optimal patient care. Current prehospital stroke triage tools range from basic stroke severity scales with limited accuracy to MSUs that can accurately diagnosis and treat patients with stroke prior to hospital arrival. Resource and technology limitations continue to hamper the growth of prehospital stroke management, but we believe several promising AI applications have the potential to address this need through greater recognition of stroke symptoms, improved dispatch, enhanced EMS training and feedback, and development of novel stroke diagnosis and triage tools.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ZW: Writing – review & editing, Writing – original draft. SE: Writing – review & editing, Conceptualization, Supervision.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Tsao CW, Aday AW, Almarzooq ZI, Anderson CAM, Arora P, Avery CL, et al. Heart disease and stroke statistics—2023 update: a report from the American Heart Association. *Circulation*. (2023) 147:E93–E621. doi: 10.1161/CIR.0000000000001123

2. Saver JL. Time is brain—quantified. *Stroke*. (2006) 37:263–6. doi: 10.1161/01.STR.0000196957.55928.ab

3. Kass-Hout T, Lee J, Tataris K, Richards CT, Markul E, Weber J, et al. Prehospital comprehensive stroke center vs primary stroke center triage in patients with suspected large vessel occlusion stroke. *JAMA Neurol*. (2021) 78:1220–7. doi: 10.1001/jamaneurol.2021.2485

4. Man S, Schold JD, Uchino K. Impact of stroke center certification on mortality after ischemic stroke: the Medicare cohort from 2009 to 2013. *Stroke*. (2017) 48:2527–33. doi: 10.1161/STROKEAHA.116.016473

5. Regenhardt RW, Mecca AP, Flavin SA, Boulouis G, Lauer A, Zachrison KS, et al. Delays in the air or ground transfer of patients for endovascular thrombectomy. *Stroke*. (2018) 49:1419–25. doi: 10.1161/STROKEAHA.118.020618

6. Stamm B, Royan R, Giurcanu M, Messe SR, Jauch EC, Prabhakaran S. Door-in-door-out times for Interhospital transfer of patients with stroke. *JAMA*. (2023) 330:636–49. doi: 10.1001/jama.2023.12739

7. Kamel H, Parikh NS, Chatterjee A, Kim LK, Saver JL, Schwamm LH, et al. Access to mechanical Thrombectomy for ischemic stroke in the United States. *Stroke*. (2021) 52:2554–61. doi: 10.1161/STROKEAHA.120.033485

8. Clawson JJ, Scott G, Gardett I, Youngquist S, Taillac P, Fivaz C, et al. Predictive ability of an emergency medical dispatch stroke diagnostic tool in identifying hospital-confirmed strokes. *J Stroke Cerebrovasc Dis*. (2016) 25:2031–42. doi: 10.1016/j.jstrokecerebrovasdis.2016.04.021

9. Heemskerk JL, Domingo RA, Tawk RG, Vivas-Buitrago TG, Huang JF, Rogers A, et al. Time is brain: prehospital emergency medical services response times for suspected stroke and effects of prehospital interventions. *Mayo Clin Proc*. (2021) 96:1446–57. doi: 10.1016/j.mayocp.2020.08.050

10. Fonarow GC, Zhao X, Smith EE, Saver JL, Reeves MJ, Bhatt DL, et al. Door-to-needle times for tissue plasminogen activator administration and clinical outcomes in acute ischemic stroke before and after a quality improvement initiative. *JAMA*. (2014) 311:1632–40. doi: 10.1001/jama.2014.3203

11. Jauch EC, Schwamm LH, Panagos PD, Barbazzeni J, Dickson R, Dunne R, et al. Recommendations for regional stroke destination plans in rural, suburban, and urban communities from the prehospital stroke system of care consensus conference: a consensus statement from the American Academy of Neurology, American Heart Association/American Stroke Association, American Society of Neuroradiology, National Association of EMS physicians, National Association of state EMS officials, society of NeuroInterventional surgery, and Society of Vascular and Interventional Neurology: Endorsed by the Neurocritical care society. *Stroke*. (2021) 52:E133–52. doi: 10.1161/STROKEAHA.120.033228

12. Nguyen TTM, van den Wijngaard IR, Bosch J, van Belle E, van Zwet EW, Dofferhoff-Vermeulen T, et al. Comparison of prehospital scales for predicting large anterior vessel occlusion in the ambulance setting. *JAMA Neurol*. (2021) 78:157–64. doi: 10.1001/jamaneurol.2020.4418

13. Grotta JC, Yamal JM, Parker SA, Rajan SS, Gonzales NR, Jones WJ, et al. Prospective, multicenter, controlled trial of Mobile stroke units. *N Engl J Med*. (2021) 385:971–81. doi: 10.1056/NEJMoa2103879

14. Ebinger M, Siegerink B, Kunz A, Wendt M, Weber JE, Schwabauer E, et al. Association between dispatch of Mobile stroke units and functional outcomes among

patients with acute ischemic stroke in Berlin. *JAMA*. (2021) 325:454–66. doi: 10.1001/jama.2020.26345

15. English SW, Barrett KM, Freeman WD, Demaerschalk BM. Telemedicine-enabled ambulances and mobile stroke units for prehospital stroke management. *J Telemed Telecare*. (2022) 28:458–63. doi: 10.1177/1357633X211047744

16. Ramos A, Guerrero WR, Pérez De La Ossa N. Prehospital Stroke Triage. *Neurology*. (2021) 97:S25–33. doi: 10.1212/WNL.0000000000012792

17. Fassbender K, Walter S, Grunwald IQ, Merzou F, Mathur S, Lesmeister M, et al. Prehospital Stroke Management in the Thrombectomy Era. *Lancet Neurol*. (2020) 19:601–10. doi: 10.1016/S1474-4422(20)30102-2

18. Bat-Erdene BO, Saver JL. Automatic acute stroke symptom detection and emergency medical systems alerting by Mobile health technologies: a review. *J Stroke Cerebrovasc Dis*. (2021) 30:105826. doi: 10.1016/j.jstrokecerebrovasdis.2021.105826

19. Wasselius J, Finn EL, Persson E, Song G, Wu S, Feng S, et al. Detection of unilateral arm paresis after stroke by wearable accelerometers and machine learning. *Sensors*. (2021) 21:7784. doi: 10.3390/s21237784

20. Walsh KB. Non-invasive sensor technology for prehospital stroke diagnosis: current status and future directions. *Int J Stroke*. (2019) 14:592–602. doi: 10.1177/1747493019866621

21. Kim HS, Kim SY, Kim YH, Park KS. A smartphone-based automatic diagnosis system for facial nerve palsy. *Sensors*. (2015) 15:26756–68. doi: 10.3390/s151026756

22. Raychev RI, Saver JL, Liebeskind DS, Penkov S, Angelov D, Stoev K, et al. Abstract WMP120: development of smartphone enabled machine learning algorithms for autonomous stroke detection. *Stroke*. (2023) 54:120. doi: 10.1161/str.54.suppl_1.WMP120

23. Schröter N, Weiller A, Rijntjes M, Harloff A, Urbach H, Kukolja J, et al. Identifying large vessel occlusion at first glance in telemedicine. *J Neurol*. (2023) 270:4318–25. doi: 10.1007/s00415-023-11775-2

24. English SW, Barrett KM, Freeman WD, Demaerschalk BM, Dumitrascu OM. Improving the telemedicine evaluation of patients with acute vision loss: a call to eyes. *Neurology*. (2022) 99:381–6. doi: 10.1212/WNL.0000000000200969

25. Pundlik S, Tomasi M, Liu R, Houston K, Luo G. Development and preliminary evaluation of a smartphone app for measuring eye alignment. *Transl Vis Sci Technol*. (2019) 8:19. doi: 10.1167/tvst.8.1.19

26. Caceres JA, Adil MM, Jadhav V, Chaudhry SA, Pawar S, Rodriguez GJ, et al. Diagnosis of stroke by emergency medical dispatchers and its impact on the prehospital care of patients. *J Stroke Cerebrovasc Dis*. (2013) 22:e610–4. doi: 10.1016/j.jstrokecerebrovasdis.2013.07.039

27. Schwartz J, Dreyer RP, Murugiah K, Ranasinghe I. Contemporary prehospital emergency medical services response times for suspected stroke in the United States. *Prehosp Emerg Care*. (2016) 20:560–5. doi: 10.3109/10903127.2016.1139219

28. Oostema JA, Chassee T, Reeves M. Emergency dispatcher stroke recognition: associations with downstream care. *Prehosp Emerg Care*. (2018) 22:466–71. doi: 10.1080/10903127.2017.1405131

29. Jia J, Band R, Abboud ME, Pajerowski W, Guo M, David G, et al. Accuracy of emergency medical services dispatcher and crew diagnosis of stroke in clinical practice. *Front Neurol*. (2017) 8:466. doi: 10.3389/fneur.2017.00466

30. Scholz ML, Collatz-Christensen H, Blomberg SNF, Boebel S, Verhoeven J, Krafft T. Artificial intelligence in emergency medical services dispatching: assessing the potential impact of an automatic speech recognition software on stroke detection taking the Capital Region of Denmark as case in point. *Scand J Trauma Resusc Emerg Med*. (2022) 30:6. doi: 10.1186/s13049-022-01020-6

31. Ceklic E, Ball S, Finn J, Brown E, Brink D, Bailey P, et al. Ambulance dispatch prioritisation for traffic crashes using machine learning: a natural language approach. *Int J Med Inform*. (2022) 168:104886. doi: 10.1016/j.ijmedinf.2022.104886

32. Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J*. (2019) 6:181–5. doi: 10.7861/fhj.2019-0036

33. Kim K, Yang H, Lee J, Lee WG. Metaverse wearables for immersive digital healthcare: a review. *Adv Sci*. (2023) 10:e2303234. doi: 10.1002/advs.202303234

34. Liaw SY, Tan JZ, Bin Rusli KD, Ratan R, Zhou W, Lim S, et al. Artificial intelligence versus human-controlled doctor in virtual reality simulation for Sepsis team training: randomized controlled study. *J Med Internet Res*. (2023) 25:e47748. doi: 10.2196/47748

35. Winkler-Schwartz A, Bissonnette V, Mirchi N, Ponnudurai N, Yilmaz R, Ledwos N, et al. Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. *J Surg Educ*. (2019) 76:1681–90. doi: 10.1016/j.jsurg.2019.05.015

36. Nogueira RG, Silva GS, Lima FO, Yeh YC, Fleming C, Branco D, et al. The FAST-ED app: a smartphone platform for the field triage of patients with stroke. *Stroke*. (2017) 48:1278–84. doi: 10.1161/STROKEAHA.116.016026

37. Turc G, Maïer B, Naggara O, Seners P, Isabel C, Tisserand M, et al. Clinical scales do not reliably identify acute ischemic stroke patients with large-artery occlusion. *Stroke*. (2016) 47:1466–72. doi: 10.1161/STROKEAHA.116.013144

38. Richard JV, Wilcock AD, Schwamm LH, Uscher-Pines L, Zachrison KS, Siddiqui A, et al. Assessment of Telestroke capacity in US hospitals. *JAMA Neurol*. (2020) 77:1035–7. doi: 10.1001/jamaneurol.2020.1274

39. Li R, St George RJ, Wang X, Lawler K, Hill E, Garg S, et al. Moving towards intelligent telemedicine: computer vision measurement of human movement. *Comput Biol Med*. (2022) 147:105776. doi: 10.1016/j.compbiomed.2022.105776

40. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. *NPJ Digit Med*. (2021) 4:5. doi: 10.1038/s41746-020-00376-2

41. Van Stigt MN, Groenendijk EA, Van Meenen LCC, van de Munckhof A, Coutinho JM, de Maat MPM, et al. Prehospital detection of large vessel occlusion stroke with EEG: results of the ELECTRA-STROKE study. *Neurology*. (2023) 101:E2522–32. doi: 10.1212/WNL.0000000000207831

42. Sense Neuro Diagnostics. Clinical Evidence for Sense Technology. (2021). Available at: https://senseneuro.com/clinical-evidence-for-sense-technology/ (Accessed 23 January 2024).

43. Kellner CP, Sauvageau E, Snyder KV, Fargen KM, Arthur AS, Turner RD, et al. The VITAL study and overall pooled analysis with the VIPS non-invasive stroke detection device. *J Neurointerv Surg*. (2018) 10:1079–84. doi: 10.1136/neurintsurg-2017-013690

44. Persson M, Fhager A, Trefna HD, Yu Y, McKelvey T, Pegenius G, et al. Microwave-based stroke diagnosis making global prehospital thrombolytic treatment possible. *IEEE Trans Biomed Eng*. (2014) 61:2806–17. doi: 10.1109/TBME.2014.2330554

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 22 May 2024
DOI 10.3389/fneur.2024.1407014

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Wei Yue,
Tianjin Huanhu Hospital, China
Adria Arboix,
Sacred Heart University Hospital, Spain

*CORRESPONDENCE
Chaohua Cui
✉ cchaiwp@163.com

RECEIVED 26 March 2024
ACCEPTED 02 May 2024
PUBLISHED 22 May 2024

CITATION
Cui C, Lan J, Lao Z, Xia T and Long T (2024) Predicting the recurrence of spontaneous intracerebral hemorrhage using a machine learning model.
Front. Neurol. 15:1407014.
doi: 10.3389/fneur.2024.1407014

COPYRIGHT
© 2024 Cui, Lan, Lao, Xia and Long. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Predicting the recurrence of spontaneous intracerebral hemorrhage using a machine learning model

Chaohua Cui*, Jiaona Lan, Zhenxian Lao, Tianyu Xia and Tonghua Long

Life Science and Clinical Medicine Research Center, Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China

Background: Recurrence can worsen conditions and increase mortality in ICH patients. Predicting the recurrence risk and preventing or treating these patients is a rational strategy to improve outcomes potentially. A machine learning model with improved performance is necessary to predict recurrence.

Methods: We collected data from ICH patients in two hospitals for our retrospective training cohort and prospective testing cohort. The outcome was the recurrence within one year. We constructed logistic regression, support vector machine (SVM), decision trees, Voting Classifier, random forest, and XGBoost models for prediction.

Results: The model included age, NIHSS score at discharge, hematoma volume at admission and discharge, PLT, AST, and CRP levels at admission, use of hypotensive drugs and history of stroke. In internal validation, logistic regression demonstrated an AUC of 0.89 and precision of 0.81, SVM showed an AUC of 0.93 and precision of 0.90, the random forest achieved an AUC of 0.95 and precision of 0.93, and XGBoost scored an AUC of 0.95 and precision of 0.92. In external validation, logistic regression achieved an AUC of 0.81 and precision of 0.79, SVM obtained an AUC of 0.87 and precision of 0.76, the random forest reached an AUC of 0.92 and precision of 0.86, and XGBoost recorded an AUC of 0.93 and precision of 0.91.

Conclusion: The machine learning models performed better in predicting ICH recurrence than traditional statistical models. The XGBoost model demonstrated the best comprehensive performance for predicting ICH recurrence in the external testing cohort.

KEYWORDS

intracerebral hemorrhage, recurrence, predicting, model, machine learning

Introduction

Patients with intracerebral hemorrhage (ICH) have higher rates of disability and mortality compared to those with ischemic stroke (1, 2). Factors such as recurrence can lead to worsening conditions and increased mortality in ICH patients (3). The recurrence rate in ICH patients is approximately 2–10% (3). Consequently, predicting the recurrence risk and

implementing preventive or therapeutic measures for patients at higher risk of recurrence is a rational strategy to improve outcomes potentially (3, 4).

Currently, there are a few studies on predicting the recurrence of ICH. One predictive model, which included 38 patients with recurrent ICH, demonstrated an AUC of 0.802 (5). However, the limited number of outcome events could lead to the model's performance instability. Additionally, the model lacked an external cohort for validation, which is crucial for assessing its generalizability. Other recurrence model had a similar condition. Therefore, a model that includes more outcome events and an external validation cohort is necessary to predict ICH recurrence more accurately.

Numerous risk factors can influence the recurrence rate (6). For example, the patient's condition, lobar cerebral hemorrhages or deep subcortical intracerebral hemorrhages, amount of bleeding, changes in the amount of bleeding, and medication status after discharge should be documented (6). These factors, with their intricate correlations, also impact the accuracy of predictions (6). All these aspects affect the performance of traditional statistical models. Machine learning models could appropriately address these challenges (7, 8). Two studies constructed machine learning models to predict outcomes in ICH patients, outperforming traditional statistical models and scores (9, 10). One study developed a machine learning model to predict the occurrence of seizures in ICH patients, which also performed well (11).

In this study, we aim to construct a machine-learning model to predict the recurrence of ICH. The model incorporates a larger cohort of patients and more outcome events. Furthermore, we validated the model's generalizability using an external cohort.

Methods

Study subjects

The training cohort of ICH patients was drawn from neurosurgery, neurology, and rehabilitation departments at the affiliated hospitals Youjiang Medical University for Nationalities. The inclusion period was from January 1, 2018, to January 1, 2021, with follow-up extending to January 1, 2022. This was a retrospective cohort. The testing cohort of ICH patients originated from the affiliated Baidong hospitals of Youjiang Medical University for Nationalities, from January 1, 2021, to January 1, 2022, and follow-up until January 1, 2023. This was a prospective cohort.

Inclusion criteria: 1. Aged 18 years or older who underwent head CT examinations and met the WHO's diagnostic criteria for ICH. 2. Standard medical or surgical treatment was administered to all patients. 3. First ICH patients.

Exclusion criteria: 1. Traumatic cerebral hemorrhage, ICH due to venous sinus thrombosis, metastatic lesions, or underlying vascular lesions. 2. Recurrence of intracerebral hemorrhage secondary to infection or hemorrhagic cerebral infarction excluded. 3. The patient died during hospitalization.

The study was conducted by the Declaration of Helsinki and the ethical standards of the institutional and/or national research committee. The study received approval from the Ethics Committee of the Affiliated Hospital of Youjiang Medical University for Nationalities. Informed consent was obtained from all study participants or their surrogates. Patients who could not give consent and those without a legally appointed representative were excluded from our study.

Risk factors and outcomes

Patient data were collected from electronic medical records, including demographic information such as age, gender, and nationality. Vital signs, including heart rate and blood pressure, were recorded at admission. Comorbidities, such as renal insufficiency, epilepsy, pneumonia, and the location of ICH, were also documented. Laboratory data, including PLT (blood platelet), INR (international normalized ratio), ALT (glutamic-pyruvic transaminase), AST (glutamic oxalacetic transaminase hemoglobin), LDL-C (low-density lipoprotein), CRP (C Reactive Protein), and others, were collected. Medical histories and medication profiles were obtained using structured questionnaires completed by patients or their relatives.

Clinical assessments, such as NIHSS (National Institute of Health stroke scale) scores at admission and discharge, mRS (Modified Rankin Scale) scores at admission and discharge, GCS (Glasgow Coma Scale) scores at admission, and ADL (activity of daily living) scores at admission, were conducted. Hematoma volumes were measured using the ABC/2 method and evaluated by the Alice software (PAREXEL International, Waltham, MA, United States) (12). Two experienced neurologists blinded to the patient's conditions and outcomes performed these evaluations at admission and discharge.

The outcome was the recurrence of ICH within 1 year post-onset. We follow up with each patient until 1 year after the onset of the disease, and document any occurrences of recurrent intracerebral hemorrhage or death within the year. Head CT examinations confirmed recurrent cases according to the WHO's ICH diagnostic criteria. The neurologists evaluated these data independently of the patient's other conditions and baseline data.

The training and testing cohorts were subjected to the same evaluation and inclusion criteria, and the uniform process collected data for both cohorts. Both cohorts shared similar baseline characteristics. All laboratory data adhered to a uniform standard for normal values.

Availability of data and material

Data from the study are available from the corresponding author upon.

Statistical analysis and machine learning model

After gathering and processing data from all patients, we performed statistical and machine learning modeling.

The number of missing features constituted less than one-fifth of the total variables in the cohort. We used multiple imputations to address missing data. Continuous variables were not converted into categorical variables. All data from the final cohort were included in the model. Statistical analyses were performed using SPSS 23.0 for

Windows and Python 3.8.0. The threshold for statistical significance was set at $p < 0.05$.

Baseline characteristics between cohorts

We presented normally distributed continuous variables as mean ± SD (standard deviation) and non-normally distributed as median and frequencies. These data were compared using a t-test for normally distributed variables (such as blood pressure, heart rate, laboratory data, etc.) or a Mann–Whitney U test for non-normally distributed variables (such as NIHSS, ADL, GCS, etc.) between the two cohorts. Categorical variables (such as gender, history of disease, history of medication, etc.) or ranked variables (such as mRS, occupation, nationality, etc.) were expressed as numbers and percentages. These data were compared using the chi-square test.

Data pre-processing

To mitigate the significant impact of class imbalance on machine learning performance, we applied random under-sampling (RUS) technique to balance the data (Scikit-learn library in Python). To guarantee fairness in the under-sampling process and unbiased generalization capabilities of the model, we performed 200 repeated experiments with distinct random under-sampling in each, coupled with ten-fold cross-validation to examine the results' stability.

Feature selection

Our data included 70 variables. We conducted feature selection using various methods to identify relevant variables for the model. We selected relevant features through Lasso regression and a step-by-step recursive procedure (Boruta library in Python). When the *p*-value of the feature is less than 0.05, the feature is included in the model. Additional features were identified and included based on clinical guidelines and literature (4–6).

Construction and internal validation of the model

We choose commonly used and well-performing basic machine learning algorithms (logistic regression, support vector machine, and decision trees) and integrated machine learning algorithms (Voting Classifier model, random forest model, and XGBoost model) that have shown good performance in previous medical research to build the model (7–11).

We constructed three fundamental machine learning algorithms: logistic regression, support vector machine (SVM), and decision trees (sklearn library in Python). We employed a five-fold cross-validation (4:1) for randomly splitting the training cohort's data. Subsequently, we performed grid search and cross-validation to optimize parameters (sklearn library in Python). We also calculated the ROC and precision values and depicted the model performance using ROC and precision plots (sklearn library in Python).

We further constructed integrated algorithm models based on three fundamental algorithms. The integrated algorithm included a Voting Classifier model, random forest model, and XGBoost model (sklearn library in Python). We calculated the F1 score for each integrated algorithm model and comprehensively compared the performance of each model through AUC, precision values, and F1 scores. We evaluated whether these integrated algorithms improved the performance of the fundamental algorithm models and selected the model with the best performance.

External validation of the model

We addressed outliers and class imbalance in the testing cohort data similarly to the training cohort data. We then predicted the recurrence of ICH in the testing cohort data (without outcome events) using the selected model from the previous step. The actual outcome events were used to validate the model's performance. We demonstrated the model's external performance and generalization ability through AUC and prediction accuracy.

Model visualization

To further elucidate the selected model, we visualized the model using feature importance and individual prediction. We depicted the importance of each feature in the final model with bar charts (SHAP library in Python). We illustrated the effect of each feature on individual patient predictions using visual representations (LIME library in Python).

Results

Baseline characteristics

Initially, 1,133 patients were eligible for the training cohort; however, we excluded 129 patients due to loss of follow-up, missing data, or withdrawal from the study. In the end, the training cohort included 1,024 patients, with 114 experiencing ICH recurrence, and 203 patients who passed away (176 due to neurological causes and 27 due to non-neurological causes). The training cohort had 53 patients with missing data, which we addressed through multiple imputations. The mean age of the training cohort was 63.50 ± 13.550 years, 43.3% of whom were female (443 patients). After excluding 24 patients for the same reasons as in the training cohort, the test cohort comprised 265 patients, with 31 experiencing ICH recurrence, and 41 patients who passed away (36 due to neurological causes and 5 due to non-neurological causes). Similarly, 11 patients in the test cohort had missing data, which we addressed through multiple imputations. The mean age of the test cohort was 59.82 ± 13.634 years, 40.4% of whom were female (107 patients) (Supplementary Figure S1).

In the training cohort, patients with recurrence were less likely to take hypotensive drugs post-onset and had less well-controlled blood pressure. Additionally, patients with recurrence were older and had greater hematoma volume at admission and discharge, higher NIHSS and ADL scores, and higher PLT levels at admission. Other features showed no significant differences (Table 1).

In the test cohort, patients with recurrence had higher NIHSS and ADL scores at discharge and greater hematoma volume at admission and discharge. Other features showed no significant differences (Table 2). Both cohorts had a similar distribution in baseline characteristics.

To compare the baseline data between the training and test cohorts, patients in the training cohort were older than those in the test cohort. Other baseline data showed no significant difference between the two cohorts (Table 3).

Data pre-processing and feature selection

Through random under-sampling and cross-validation, the model performance has improved by approximately 10%.

Through Lasso regression and a stepwise process, we identified several relevant variables for the model, including age, NIHSS score at

TABLE 1 Baseline characteristic of training cohort.

Feature	All patients (*N* = 1,024)	Recurrence (*N* = 114)	No-Recurrence (*N* = 910)	*p**
Age, years	63.50 (13.550)	68.81 (12.960)	62.84 (13.483)	<0.001
Female, %	443 (43.3)	48 (42.1)	395 (43.4)	0.361
Discharge NIHSS score	6 (1–15)	20 (15–35)	4 (1–11)	<0.001
Admission ADL score	40 (20–70)	25 (10–45)	45 (25–75)	<0.001
Admission hematoma, mL	21 (11–24)	25 (22–28)	19 (10–24)	<0.001
Discharge hematoma, mL	13 (7–22)	36 (32–40)	12 (6–19)	<0.001
History of stroke, %	277 (27.1)	26 (22.8)	251 (27.6)	0.137
Hypotensive drugs, %	826 (80.7)	67 (58.8)	759 (83.4)	0.003
Platelet, mmol/L	202.80 (71.176)	179.78 (64.950)	205.69 (71.429)	<0.001
AST, mmol/L	26.48 (16.340)	26.57 (17.264)	26.47 (16.230)	0.949
CRP, mmol/L	19.55 (16.975)	20.88 (17.470)	19.39 (17.155)	0.375

*p** was calculated by ANOVA, Chi-square test, or Mann–Whitney U test as appropriate. NIHSS, National Institute of Health stroke scale; ADL, activity of daily living; AST, glutamic oxalacetic transaminase; CRP, C-reactive protein.

TABLE 2 Baseline characteristic of testing cohort.

Feature	All patients (*N* = 265)	Recurrence (*N* = 31)	No-Recurrence (*N* = 234)	*p**
Age, years	59.82 (13.634)	61.48 (12.951)	59.60 (13.733)	0.471
Female, %	107 (40.4)	13 (41.9)	94 (40.2)	0.615
Discharge NIHSS score	5 (1–17)	18 (11–35)	4 (1–12)	<0.001
Admission ADL score	50 (0–85)	45 (10–65)	65 (15–90)	<0.001
Admission hematoma, mL	18 (11–25)	28 (26–29)	16 (10–23)	<0.001
Discharge hematoma, mL	12 (6–25)	30 (27–32)	10 (5–20)	<0.001
History of stroke, %	56 (21.1)	6 (19.4)	50 (21.4)	0.657
Hypotensive drugs, %	228 (86.0)	24 (77.4)	204 (87.2)	0.198
Platelet, mmol/L	200.39 (70.061)	193.45 (52.069)	202.19 (72.113)	0.444
AST, mmol/L	26.77 (15.174)	27.32 (18.707)	26.70 (14.689)	0.831
CRP, mmol/L	19.67 (14.017)	19.93 (13.336)	19.62 (14.132)	0.910

*p** was calculated by ANOVA, Chi-square test, or Mann–Whitney U test as appropriate. NIHSS, National Institute of Health stroke scale; ADL, activity of daily living; AST, glutamic oxalacetic transaminase; CRP, C-reactive protein.

TABLE 3 Baseline characteristic between training and testing cohort.

Feature	Training set (*N* = 1,024)	Testing set (*N* = 265)	*p**
Age, years	63.50 (13.550)	59.82 (13.634)	<0.001
Female, %	443 (43.3)	107 (40.4)	0.317
Discharge NIHSS score	6 (1–15)	5 (1–17)	0.279
Admission ADL score	40 (20–70)	50 (0–85)	0.273
Admission hematoma, mL	21 (11–24)	18 (11–25)	0.116
Discharge hematoma, mL	13 (7–22)	12 (6–25)	0.907
History of stroke, %	277 (27.1)	56 (21.1)	0.317
Hypotensive drugs, %	826 (80.7)	228 (86.0)	0.203
Platelet, mmol/L	202.80 (71.176)	200.39 (70.061)	0.622
AST, mmol/L	26.48 (16.340)	26.77 (15.174)	0.790
CRP, mmol/L	19.55 (16.975)	19.67 (14.017)	0.917

*p** was calculated by ANOVA, Chi-square test, or Mann–Whitney U test as appropriate. NIHSS, National Institute of Health stroke scale; ADL, activity of daily living; AST, glutamic oxalacetic transaminase; CRP, C-reactive protein.

discharge, hematoma volume at admission and discharge, and PLT, AST, and CRP levels at admission. Subsequently, based on clinical significance and literature (4–6), and for enhanced model performance after adding additional variables, particularly the improvement in model performance during external validation, we have chosen to incorporate the use of hypotensive drugs and history of stroke into the model.

Internal validation performance of models

For the basic machine learning models, internal validation revealed that the AUC for the logistic regression model was 0.89, and the precision was 0.81, the positive predictive value (PPV) was 87.3%, and negative predictive value (NPV) was 76.0%. The SVM model had an AUC of 0.93 and a precision of 0.90, a PPV of 94.2% and a NPV of 88.5%, while the decision trees model showed an AUC of 0.91 and a precision of 0.71, a PPV of 79.0% and a NPV of 81.5%. All three models demonstrated excellent discrimination, with the logistic regression and SVM models also exhibiting excellent calibration.

For the integrated machine learning models, the score for the Logistic Regression model was 0.951, for the Decision Tree Classifier was 0.962, and for the SVC model was 0.883. In the Voting Classifier model, the score for the soft voting method was 0.928, with an AUC and precision of 0.96 and 0.86, a PPV of 94.2% and a NPV of 89.4%, respectively. The random forest model scored 0.966 and exhibited an AUC and precision of 0.95 and 0.93, a PPV of 95.0% and a NPV of 92.3%, respectively. The XGBoost model scored 0.977, with an AUC and precision of 0.95 and 0.92, a PPV of 94.6% and a NPV of 93.1%, respectively. All three integrated machine learning models exhibited excellent discrimination, and the latter two models demonstrated excellent calibration. The random forest and XGBoost algorithms improved the basic models' performance.

External validation performance of models

External validation for the basic machine learning models showed that the AUC for the logistic regression model was 0.81 (Supplementary Figure S3a-1), with a precision of 0.79 (Supplementary Figure S3a-2). The SVM model had an AUC of 0.87 (Supplementary Figure S3b-1) and a precision of 0.76 (Supplementary Figure S3b-2), and the decision trees model had an AUC of 0.71 (Supplementary Figure S3c-1) and a precision of 0.70 (Supplementary Figure S3c-2).

For the integrated machine learning models, the random forest model exhibited an AUC of 0.92 (Supplementary Figure S4a-1) and a precision of 0.86 (Supplementary Figure S4a-2), while the XGBoost model had an AUC of 0.93 (Supplementary Figure S4b-2) and a precision of 0.91 (Supplementary Figure S4b-2). The random forest model had a PPV of 90.0% and a NPV of 85.2%, and the XGBoost model had a PPV of 92.5% and a NPV of 93.2%. The integrated machine learning algorithms still demonstrated excellent discrimination and calibration in the test cohort.

Model visualization

In the feature importance figure of the XGBoost model, we found that the value of the hematoma and NIHSS scores at discharge had the most significant effect on the model. Then, the age, the value of PLT and AST, and the value of hematoma at admission also affected the model differently (Figure 1). In the feature importance figure of the random forest model, we found that the value of hematoma at discharge had the best significant effect on the model. Then, the NIHSS score at discharge, the value of hematoma at admission, age, and the value of PLT and AST also affected the model differently (Supplementary Figure S2).

Figure 2 shows the process of the XGBoost model to predict the recurrence rate in individual ICH patients. First, the XGBoost model calculated the score through the difference in volume extent of hematoma and NIHSS score, age, PLT, and AST. Then, the model calculated the possibility of ICH recurrence in the individual patients. For the selected patients, the model accurately predicts the patients would have recurrent ICH with a 100% possibility.

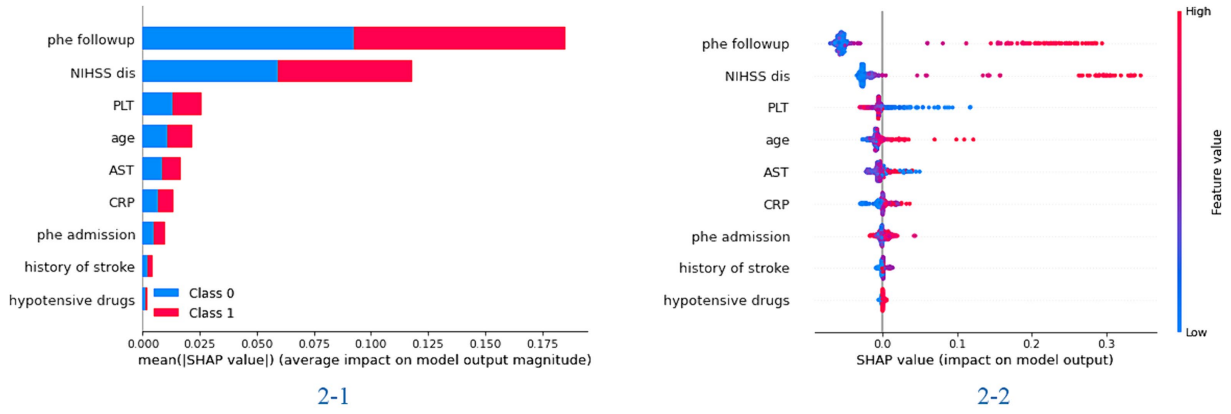


FIGURE 1
The importance of the feature plot of the XGBoost model. 2-1: the importance of features for predicting in the total model; 2-2: the importance of features for predicting in each patient of the model.

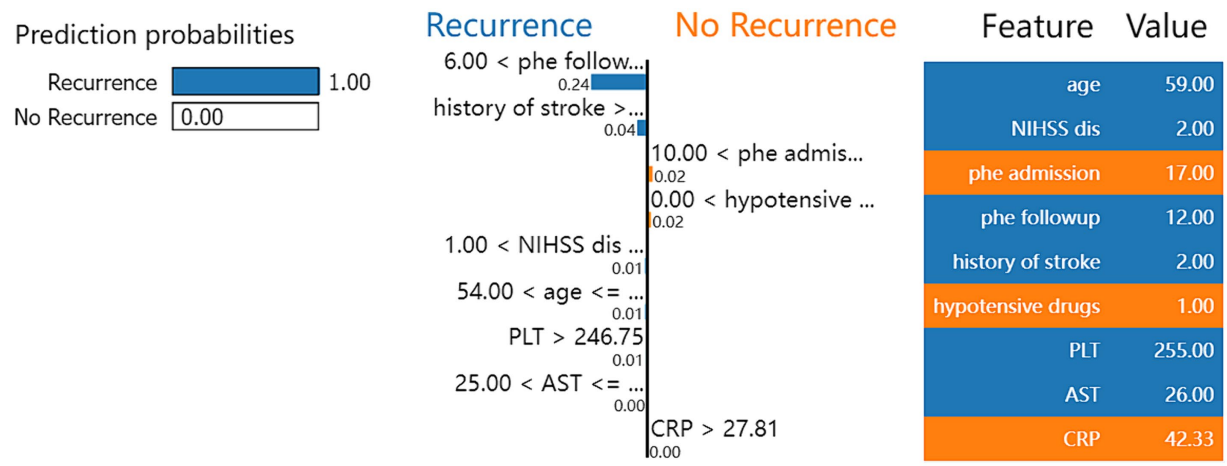


FIGURE 2
Local interpretable model-agnostic explanation (LIME) with XGBoost model applied to one correctly predicted patient that ICH recurrence within 1 year.

Discussion

The study constructs machine learning models to predict the recurrence of ICH. We evaluated six machine learning models. The models included age, NIHSS score at discharge, and hematoma volume at admission and discharge. The SVM and decision trees were performed better than logistic regression model in the training cohort. The integrated machine learning models performed better than the basic models in the training cohort. In the test cohort, the AUC and precision of the basic models all decreased to less than 0.9. However, the integrated models, such as the random forest and XGBoost models, still performed better. Notably, the XGBoost model performed excellently in the testing cohort, with an AUC and precision greater than 0.9.

We performed multiple imputation on data with a small number of missing features, and data with more than 30% missing features were not included in the analysis. Less than 1% of the data was not included in the analysis, and it did not significantly affect the result analysis. We standardized all the data included in our model; thus, the data distribution did not significantly affect the results.

For selected features, the NIHSS score at discharge had a highly significant impact on the model. Their importance was ranked second and third in the final random forest and XGBoost models. The neurological deficit at discharge contained more information for predicting long-term events. The discharge score at least reflects the illness's initial state and therapy's effect. In other study, the discharge score had a more pronounced relationship with patient outcomes (5, 9–11). The hematoma volume also had a similar effect. Our model included hematoma volume at admission and discharge, allowing it to evaluate the changing situation throughout hospitalization. These changes could more accurately reflect the patients' conditions. This was similar to other models, including the GCS score at admission and discharge (5). The hematoma volume consistently ranked among the top three in our final random forest and XGBoost models, significantly improving our model's performance.

Our model also included age and PLT levels. It is plausible that older age and lower PLT levels could increase the risk of ICH recurrence, which is consistent with other studies (13). Our model also indicated that AST and CRP levels affect the risk of recurrence. Poor liver function and infection can affect the state of the illness (14, 15). However, these effects require further investigation. Our model includes a history of stroke and the use of hypotensive drugs. Although these two features did not significantly affect our model performance, these clinically significant features could improve the model's generalizability. Therefore, our model could perform better in different region cohorts. An acute intraparenchymal hematoma in a young patient can be the presenting manifestation of a hematological disorder (16). The hematologic disorders from other stroke etiologies that have a different treatment approach and outcome (16). For less young patients in our cohort, a model included more young patients was more suit for these patients with hematological disorder.

The recurrence of ICH can be influenced by numerous risk factors that vary over time. These factors exhibit a complex interrelationship that traditional statistical models struggle to represent. Machine learning models are more adept at processing these complex characteristics. Consequently, machine learning models have the potential to identify novel predictors with enhanced predictive value. They demonstrated excellent performance in the training cohort due to the abovementioned advantages. However, machine learning models face challenges when applied to external datasets. In other words, they may exhibit poorer generalization performance. The overall performance of our model decreased in the test cohort. Nevertheless, the integrated model exhibited superior generalization capabilities, maintaining its excellent performance even within the test cohort. Consequently, the integrated model could hold significant clinical value across various regions.

Another common challenge with machine learning models is the interpretability of their outcomes. The SHAP library is widely used for global interpretation of model outcomes, revealing the significance of each model feature (17). This means the library elucidates the impact and contribution of each feature to the predictive results and displays each feature's value within the model. The LIME library, which stands

for Local Interpretable Model-Agnostic Explanations, can demonstrate the model's prediction process for individual samples (18). Consequently, LIME visually illustrates how the model predicts the likelihood of outcome events based on the value of each feature. These two libraries enhance our model's interpretability, demonstrating practical value and superior performance in clinical settings.

This study faced several limitations. First, the training cohort might be susceptible to bias owing to its retrospective design. Second, the training and testing cohorts, which originated from the same geographical region, had a similar feature distribution. Consequently, it becomes imperative to validate the model's performance across diverse regional patient populations further. Third, our model did not incorporate neuroimaging data; including a broader array of data types could potentially enhance model performance.

Conclusion

The machine learning model demonstrated excellent performance in predicting the recurrence of ICH compared to traditional statistical models. The XGBoost models with hematoma volume at admission and discharge outperformed the external test cohort. The change situation in hematoma volume improves the performance of model. The XGBoost model exhibited the best discrimination and calibration in predicting the recurrence of ICH. The XGBoost model offered excellent predictive value for guiding medical care in patients at a higher risk of ICH recurrence.

The use of this XGBoost model, along with other machine learning models for predicting recurrent intracerebral hemorrhage, can improve the accuracy of recurrent intracerebral hemorrhage prediction. It guides the formulation of personalized treatment and prevention plans in clinical medicine. The next step involves validating the model's applicability in other regions and ethnicities to confirm its practicality and universality.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Affiliated Hospital of Youjiang Medical University for Nationalities. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CC: Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Supervision, Writing – original draft, Writing – review & editing. JL: Conceptualization, Formal analysis, Writing – review & editing. ZL: Data curation, Investigation, Methodology, Software, Writing – review & editing. TX: Data curation, Investigation, Methodology, Writing – review & editing. TL: Data curation, Investigation, Software, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Guangxi Zhuang Autonomous Region Health and Family Planning Commission (Z-B20221499), Liuzhou Scientific Research Technological Development Programs (2022CAC0118), and Specific Research Project of Guangxi for Research Bases and Talents (AD23026241).

Acknowledgments

We are grateful to MD Lijun Lao for providing valuable suggestions and guidance for the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1407014/full#supplementary-material

References

1. Hemphill JC, Greenberg SM, Anderson CS, Becker K, Bendok BR, Cushman M, et al. Guidelines for the management of spontaneous intracerebral hemorrhage. *Stroke*. (2015) 46:2032–60. doi: 10.1161/str.0000000000000069

2. Geurts M, Macleod MR, van Thiel G, van Gijn J, Kappelle LJ, van der Worp HB. End-of-life decisions in patients with severe acute brain injury. *Lancet Neurol*. (2014) 13:515–24. doi: 10.1016/s1474-4422(14)70030-4

3. Wolf ME, Alonso A, Ebert AD, Szabo K, Chatzikonstantinou A. Etiologic and clinical characterization of patients with recurrent spontaneous intracerebral hemorrhage. *Eur Neurol*. (2016) 76:295–301. doi: 10.1159/000452659

4. Huhtakangas J, Löppönen P, Tetri S, Juvela S, Saloheimo P, Bode MK, et al. Predictors for recurrent primary intracerebral hemorrhage a retrospective population-based study. *Stroke*. (2013) 44:585–90. doi: 10.1161/strokeaha.112.671230

5. Zhang S, Zhang X, Ling Y, Li A. Predicting recurrent hypertensive intracerebral hemorrhage: derivation and validation of a risk-scoring model based on clinical characteristics. *World Neurosurg*. (2019) 127:E162–71. doi: 10.1016/j.wneu.2019.03.024

6. de Mendiola J, Arboix A, García-Eroles L, Sánchez-López MJ. Acute spontaneous lobar cerebral hemorrhages present a different clinical profile and a more severe early prognosis than deep subcortical intracerebral hemorrhages-a hospital-based stroke registry study. *Biomedicines*. (2023) 11:11. doi: 10.3390/biomedicines11010223

7. Cui CH, Li YC, Liu SH, Wang P, Huang Z. The unsupervised machine learning to analyze the use strategy of statins for ischaemic stroke patients with elevated transaminase. *Clin Neurol Neurosurg*. (2023) 232:107900. doi: 10.1016/j.clineuro.2023.107900

8. Cui CH, Li CH, Hou M, Wang P, Huang Z. The machine learning methods to analyze the using strategy of antiplatelet drugs in ischaemic stroke patients with gastrointestinal haemorrhage. *BMC Neurol*. (2023) 23:369. doi: 10.1186/s12883-023-03422-0

9. Wang HL, Hsu WY, Lee MH, Weng HH, Chang SW, Yang JT, et al. Automatic machine-learning-based outcome prediction in patients with primary intracerebral hemorrhage. *Front Neurol*. (2019) 10:7. doi: 10.3389/fneur.2019.00910

10. Hall AN, Weaver B, Liotta E, Maas MB, Faigle R, Mroczek DK, et al. Identifying modifiable predictors of patient outcomes after intracerebral hemorrhage with machine learning. *Neurocrit Care*. (2021) 34:73–84. doi: 10.1007/s12028-020-00982-8

11. Bunney G, Murphy J, Colton K, Wang H, Shin HJ, Faigle R, et al. Predicting early seizures after intracerebral hemorrhage with machine learning. *Neurocrit Care*. (2022) 37:322–7. doi: 10.1007/s12028-022-01470-x

12. Kothari U, Brott T, Broderick JP, Barsan WG, Sauerbeck LR, Zuccarello M, et al. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke*. (1996) 27:1304–5. doi: 10.1161/01.Str.27.8.1304

13. Naidech AM, Bendok BR, Garg RK, Bernstein RA, Alberts MJ, Bleck TP. Reduced platelet activity is associated with more intraventricular hemorrhage. *Neurology*. (2009) 72:A73–3. doi: 10.1227/01.NEU.0000351769.39990.16

14. Ruban A, Daya N, Schneider ALC, Gottesman R, Selvin E, Coresh J, et al. Liver enzymes and risk of stroke: the atherosclerosis risk in communities (ARIC) study. *J Stroke*. (2020) 22:357. doi: 10.5853/jos.2020.00290

15. Diedler J, Sykora M, Hahn P, Rupp A, Rocco A, Herweh C, et al. C-reactive-protein levels associated with infection predict short- and long-term outcome after Supratentorial intracerebral hemorrhage. *Cerebrovasc Dis*. (2009) 27:272–9. doi: 10.1159/000199465

16. Arboix A, Besses C. Cerebrovascular disease as the initial clinical presentation of haematological disorders. *Eur Neurol*. (2004) 37:207–11. doi: 10.1159/000117444

17. Wang K, Tian J, Zheng C, Yang H, Ren J, Liu Y, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. *Comput Biol Med*. (2021) 137:9. doi: 10.1016/j.compbiomed.2021.104813

18. Hussain I, Jany R, Boyer R, Azad AKM, Alyami SA, Park SJ, et al. An explainable EEG-based human activity recognition model using machine-learning approach and LIME. *Sensors*. (2023) 23:15. doi: 10.3390/s23177452

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 22 May 2024
DOI 10.3389/fneur.2024.1368902

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Yingwei Guo,
Northeast Petroleum University, China
Xueli Chang,
First Hospital of Shanxi Medical University, China

*CORRESPONDENCE
Longbin Jia
✉ sxjcjlb@163.com
Lina Xu
✉ sxjcxln@163.com
Fengbing Yang
✉ yangfb8008@163.com
Zixuan Guo
✉ 109291035@qq.com

RECEIVED 11 January 2024
ACCEPTED 24 April 2024
PUBLISHED 22 May 2024

CITATION
Liu W, Jia L, Xu L, Yang F, Guo Z, Li J, Zhang D, Liu Y, Xiang H, Cheng H, Hou J, Li S and Li H (2024) Prediction of early neurologic deterioration in patients with perforating artery territory infarction using machine learning: a retrospective study. *Front. Neurol.* 15:1368902. doi: 10.3389/fneur.2024.1368902

COPYRIGHT
© 2024 Liu, Jia, Xu, Yang, Guo, Li, Zhang, Liu, Xiang, Cheng, Hou, Li and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Prediction of early neurologic deterioration in patients with perforating artery territory infarction using machine learning: a retrospective study

Wei Liu[1], Longbin Jia[1]*, Lina Xu[1]*, Fengbing Yang[1]*, Zixuan Guo[1]*, Jinna Li[1], Dandan Zhang[1], Yan Liu[2], Han Xiang[2], Hongjiang Cheng[1], Jing Hou[1], Shifang Li[1] and Huimin Li[1]

[1]Department of Neurology, Jincheng People's Hospital, Jincheng, China, [2]The First Clinical College of Changzhi Medical College, Changzhi, China

Background: Early neurological deterioration (END) is a frequent complication in patients with perforating artery territory infarction (PAI), leading to poorer outcomes. Therefore, we aimed to apply machine learning (ML) algorithms to predict the occurrence of END in PAI and investigate related risk factors.

Methods: This retrospective study analyzed a cohort of PAI patients, excluding those with severe stenosis of the parent artery. We included demographic characteristics, clinical features, laboratory data, and imaging variables. Recursive feature elimination with cross-validation (RFECV) was performed to identify critical features. Seven ML algorithms, namely logistic regression, random forest, adaptive boosting, gradient boosting decision tree, histogram-based gradient boosting, extreme gradient boosting, and category boosting, were developed to predict END in PAI patients using these critical features. We compared the accuracy of these models in predicting outcomes. Additionally, SHapley Additive exPlanations (SHAP) values were introduced to interpret the optimal model and assess the significance of input features.

Results: The study enrolled 1,020 PAI patients with a mean age of 60.46 (range 49.11–71.81) years. Of these, 30.39% were women, and 129 (12.65%) experienced END. RFECV selected 13 critical features, including blood urea nitrogen (BUN), total cholesterol (TC), low-density-lipoprotein cholesterol (LDL-C), apolipoprotein B (apoB), atrial fibrillation, loading dual antiplatelet therapy (DAPT), single antiplatelet therapy (SAPT), argatroban, the basal ganglia, the thalamus, the posterior choroidal arteries, maximal axial infarct diameter (measured at $<$15 mm), and stroke subtype. The gradient-boosting decision tree had the highest area under the curve (0.914) among the seven ML algorithms. The SHAP analysis identified apoB as the most significant variable for END.

Conclusion: Our results suggest that ML algorithms, especially the gradient-boosting decision tree, are effective in predicting the occurrence of END in PAI patients.

KEYWORDS

early neurologic deterioration, perforating artery territory infarction, machine learning, prediction, apolipoprotein B

1 Introduction

Perforating artery territory infarction (PAI), a subtype of single subcortical infarction (SSI) caused by the occlusion of a perforating artery, is frequently observed in acute ischemic stroke, accounting for approximately 15.3–25% of all stroke cases (1, 2). The etiology of PAI may involve several mechanisms, such as lipohyalinosis, large plaques in the parent artery, and microatheroma (3). Lipohyalinosis is a vasculopathy that affects cerebral small vessels, leading to a "lacunar infarct" (LI), a major contributor to PAI (4, 5). Similarly, large plaques in the parent artery with severe stenosis may result in perforating artery embolisms. Additionally, microatheromas may give rise to branch atheromatous disease (BAD), characterized by ischemic lesions of ≥15 mm in diameter, typically observed on radiological imaging and the absence of severe stenosis of the parent artery (6).

Early neurological deterioration (END), characterized as a rapid exacerbation of neurological symptoms during the acute phase of a stroke, has been observed in 20–43% of PAI patients (7–9) and is linked to unfavorable patient outcomes (10, 11). The efficacy of treatment strategies for END in PAI patients, particularly those without severe stenosis of the parent artery, such as in LI and BAD cases, remains uncertain. The prediction of END in these patients is complex and heterogeneous, posing challenges to clinical management. Thus, identifying risk factors, pinpointing high-risk patients, and implementing timely interventions are essential for managing END in PAI patients without severe stenosis of the parent artery.

Advancements in computing power, the proliferation of big data, and the evolution of algorithms have significantly propelled machine learning (ML) in disease prediction (12). ML algorithms, a crucial aspect of artificial intelligence, excel at discerning patterns within intricate datasets using computational methods (13, 14). Compared to traditional statistics, ML shows greater proficiency in forecasting complex clinical events influenced by numerous factors and variables (15). Consequently, this research aims to develop ML models adept at predicting END in PAI patients without severe stenosis of the parent artery, utilizing data from a real-world, single-center cohort database.

2 Methods

2.1 Study design and patients

This retrospective, observational study was conducted on a cohort of PAI patients at Jincheng People's Hospital from September 2016 to July 2022. The inclusion criteria were as follows: patients aged 18 years or older, those diagnosed with PAI, and those admitted within 24 h of symptom onset. PAI was characterized as a single, small subcortical infarction in the territory of a perforating arteriole (16), identified by magnetic resonance imaging (MRI), and without significant large vessel stenosis (>50%), as confirmed by magnetic resonance angiography (MRA) or computed tomography angiography (CTA), with no maximum diameter limit (17). The exclusion criteria encompassed patients with multiple or cortical lesions, a premorbid modified Rankin Scale (mRS) score of ≥2, stroke mimics, or MRI-negative stroke. The study was approved by the ethics committee of Jincheng People's Hospital, and written informed consent was waived due to its retrospective nature; all patient information was anonymized before analysis.

2.2 Baseline data

We collected baseline data, including demographic details such as age, sex, and body mass index (BMI). We also gathered information on current smoking and drinking habits (≥20 g/day), medical history (including stroke, hypertension, diabetes, coronary atherosclerotic heart disease, and atrial fibrillation), secondary prevention treatment, laboratory data, clinically significant features [such as time from onset to presentation, National Institutes of Health Stroke Scale (NIHSS) score, pre-stroke mRS score], acute phase treatment [including IV thrombolysis with alteplase, loading dose dual antiplatelet therapy (DAPT, 100 mg aspirin and 300 mg clopidogrel), single antiplatelet therapy (SAPT), lipid-lowering drugs, and argatroban], and radiological characteristics.

Radiological characteristics encompassed location (such as the internal capsule, basal ganglia, thalamus, pons, lateral ventricle, and centrum semiovale), the culprit vessel supplying the basal ganglia (including the lenticulostriate artery (LSA), posterior choroidal artery, and the recurrent artery of Heubner), stroke subtypes (such as LI and BAD), maximum axial infarct diameter, layers of cerebral infarct lesions, and white matter hyperintensities (WMH). The definition of BAD is based on infarct lesions observed in transversal diffusion-weighted imaging (DWI) scans that extend for at least three consecutive slices within LSA terminations or unilateral involvement of the pons connected to the cerebral surface of the ventral pons without crossing the midline in the paramedian pontine artery (PPA) terminations (1, 18). Maximal axial infarct diameter and layers of cerebral infarct lesions were measured in transversal DWI scans at baseline, with WMH severity at baseline categorized using the modified Fazekas scale (19).

Following data collection, a total of 65 variables were included in the baseline data analysis.

2.3 Outcome definition

The primary outcome of this analysis was END, defined as an increase in the NIHSS score of ≥2 and a rise in the motor component of the NIHSS score of ≥1 compared to the initial NIHSS score within 7 days of hospital admission.

2.4 Machine learning algorithms

In this study, we utilized seven ML models to predict END in PAI patients: logistic regression (LR) (20), random forest (RF) (21), adaptive boosting (AdaBoost) (22), gradient boosting decision tree (GBDT) (23), histogram-based gradient boosting (HGB) (24, 25), extreme gradient boosting (XGBoost) (26), and category boosting (CatBoost) (27, 28). The best-performing model was selected based on its evaluation metrics. LR, suited for binary classification problems, predicts outcomes by converting a linear

function into a sigmoid function, ranging from 0 to 1. RF, an ensemble learning method, combines multiple decision trees, each built from randomly selected subsets of training data and features, to enhance performance and generalizability. AdaBoost, GBDT, HGB, XGBoost, and CatBoost are ensemble learning methods that strengthen models by sequentially training multiple weak learners, with each new model focusing on correcting the errors of its predecessors. The following provides a comprehensive examination of the five ensemble learning methodologies.

2.4.1 AdaBoost

The fundamental principle of the AdaBoost algorithm involves categorizing a collection of weak learners through a process of weighted majority voting, or summation. This method takes into account the errors committed by preceding weak learners and consistently refines the dataset (22, 29).

2.4.2 GBDT

The core concept of GBDT involves employing a gradient-boosting methodology to train a series of decision trees. In each iterative training phase, GBDT computes the residual or gradient of the existing model and utilizes it as the training objective for the subsequent decision tree. The incorporation of new decision trees aims to approximate this residual or gradient, thereby progressively enhancing the performance of the model (23).

2.4.3 HGB

HGB refers to the implementation of gradient boosting algorithms, particularly the popular XGBoost and LightGBM libraries, which use histograms to approximate the distribution of the data. This approach enhances the efficiency and scalability of the model training process without sacrificing much accuracy (24, 25).

XGBoost, a sophisticated and scalable machine learning technique, is renowned for its exceptional proficiency in efficiently managing missing data and seamlessly integrating weak predictive models to form a more precise one. It employs a second-order Taylor expansion to compute the loss function, thereby exhibiting superior performance in both computational speed and prediction accuracy (30, 31).

2.4.4 CatBoost

A machine learning library that has been designed to efficiently handle categorical features. Developed by Yandex, it is renowned for its superior performance in gradient-boosting algorithms, particularly when applied to datasets that contain both numerical and categorical variables (28).

2.5 Data processing

In this study, continuous variables were imputed using the median values for each variable to address missing values. Categorical variables were converted into numerical values through dummy encoding. All numerical values were then standardized to ensure uniformity in scale and precision in comparisons.

2.6 Feature selection

Feature selection (32), a process used to eliminate superfluous features from a large dataset, improves a machine learning model's efficiency. We employed recursive feature elimination with cross-validation (RFECV), a prominent algorithm in feature selection, which methodically removes the least important features to pinpoint the most effective subset. This study used RFECV based on logistic regression for optimal variable selection.

2.7 Model derivation and validation

Patients were randomly divided into training and testing sets at a 7:3 ratio. For model derivation, a shuffle-split cross-validation method was employed to prevent overfitting to a specific dataset and to ensure model generalizability. Shuffle-split cross-validation (33) is a resampling technique employed in machine learning to assess model performance on a constrained data sample. This method entails the random division of the dataset into two subsets: one designated for training and the other for testing. The procedure is executed multiple times, or "folds", to yield an average performance metric. Widely adopted in machine learning, shuffle-split cross-validation ensures that model efficacy remains robust, avoiding overreliance on specific data partitioning. This approach effectively reduces biases and offers a more dependable estimation of the model's generalization capabilities for novel, unseen datasets.

Additionally, GridSearch CV with shuffle-split cross-validation was utilized to fine-tune and optimize the model hyperparameters on the training set. Supplementary Table 1 details the selected parameter values for each algorithm in the grid-search process. After optimization on the training set, model performance in the testing set was assessed using various metrics, including receiver operating characteristic (ROC) curve, accuracy, F1-score, Matthew's correlation coefficient (MCC), specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), and Youden's index.

2.8 Model interpretation

To evaluate the importance of each variable, the Shapley Additive exPlanations (SHAP) values were utilized to interpret the machine learning model. Originating from cooperative game theory (34), SHAP assigns an importance value to each feature for a given prediction. A positive SHAP value signifies a beneficial impact on the model's prediction, whereas a negative value indicates an adverse effect. The SHAP method thus serves as a vital tool for understanding and interpreting the behavior of machine learning models.

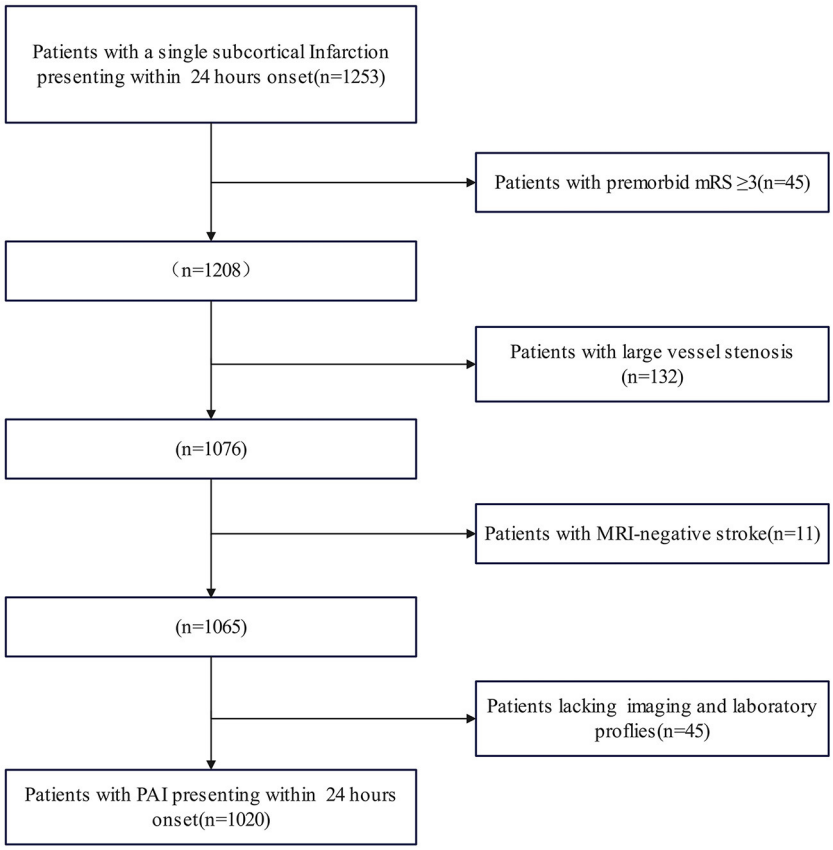


FIGURE 1
Flow diagram of the included patients.

2.9 Statistical analysis

Patients were divided into two groups based on the END outcome: the END group and the clinically stable group. Continuous variables were presented as mean ± standard deviation for normally distributed variables and as medians with interquartile ranges (IQRs) for non-normally distributed data. Categorical variables were expressed as percentages. Categorical variables were analyzed using Fisher's exact test or a $\chi 2$ test, while continuous variables were assessed using the Student's *t*-test or Mann–Whitney *U*-test. Statistical analyses were conducted using IBM SPSS Statistics for Windows, version 25.0 software (IBM Corp., Armonk, NY, USA), and a two-sided *p*-value of ≤ 0.05 was considered statistically significant. The ML algorithms were implemented using Python software (version 3.9).

3 Results

3.1 Baseline characteristics

The study initially included 1,273 patients, but 1,020 PAI patients were enrolled for evaluation after excluding 253 subjects due to missing data and fulfilling the exclusion criteria. Figure 1 presents the patient flow diagram. The baseline characteristics are detailed in Table 1. The average age of the 1,020 patients was 60.46 (range 49.11–71.81) years, with 30.39% of them being women. The END group comprised 129 (12.65%) patients with an average age of 59.08 (range 48.29–69.87) years, and 27.91% were women; the clinically stable group included 891 (87.35%) patients, averaging 60.66 (range 49.24–72.08) years, with 30.74% being women. The median time to END onset was 16 (range 5–25) h. The univariate analysis indicated that factors such as apolipoprotein B (Apo B), the ratio of apolipoprotein A1 (ApoA1) to ApoB, admission NIHSS score, motor arm NIHSS score, motor leg NIHSS score, facial palsy NIHSS score, alteplase, argatroban, lesion location (including the basal ganglia, internal capsule, lateral ventricle, and thalamus), lenticulostriate artery, stroke subtype, and maximal axial infarct diameter were significantly associated with an increased risk of END.

3.2 Features selection

Based on the RFECV results, 13 variables were identified for the predictive model, including blood urea nitrogen (BUN), total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), apolipoprotein B (apoB), atrial fibrillation, loading-dose dual antiplatelet therapy (DAPT), single antiplatelet therapy (SAPT), argatroban, the basal ganglia, the thalamus, the posterior choroidal

10.3389/fneur.2024.1368902

TABLE 1 Baseline variables for the total population and for both groups.

Variables	Clinically stable ($n = 891$)	END ($n = 129$)	Total ($n = 1,020$)	*P*-value
Demographics				
Women, *n* (%)	274 (30.74)	36 (27.91)	310 (30.39)	0.5114
Age, mean (SD)	60.66 ± 11.42	59.08 ± 10.79	60.46 ± 11.35	0.1406
BMI, kg/m^2 (IQR)	24.22 (22.49–26.66)	24.21 (22.82–26.34)	24.22 (22.55–26.66)	0.698
Medical history *n* (%)				
Hypertension	665 (74.63)	97 (75.19)	762 (74.71)	0.8915
Diabetes	165 (18.52)	29 (22.48)	194 (19.02)	0.2839
Coronary heart disease	65 (7.30)	6 (4.65)	71 (6.96)	0.2701
Atrial fibrillation	13 (1.46)	0 (0.00)	13 (1.27)	0.1674
Ischemic stroke	148 (16.61)	16 (12.4)	164 (16.08)	0.224
Hemorrhagic stroke	28 (3.14)	3 (2.32)	31 (3.04)	0.8632
Secondary prevention treatment, *n* (%)				
Antihypertensive treatment	390 (43.77)	46 (35.66)	436 (42.75)	0.0817
Antidiabetic treatment	109 (12.23)	19 (14.73)	128 (12.55)	0.424
Lipid-lowering treatment	150 (16.84)	18 (13.95)	168 (16.47)	0.4069
Antiplatelet therapy	155 (17.40)	21 (16.28)	176 (17.25)	0.7536
Anticoagulant therapy	6 (0.67)	0 (0.00)	6 (0.59)	0.3499
Smoking history, *n* (%)	202 (22.67)	29 (22.48)	231 (22.65)	0.9615
Drinking history (≥20 g/day), *n* (%)	74 (8.31)	13 (10.08)	87 (8.53)	0.5006
Laboratory data				
WBC count, *10^9 (IQR)	6.40 (5.30–7.70)	6.70 (5.49–7.85)	6.43 (5.32–7.70)	0.3883
RBC count, *10^{12} (SD)	4.71 ± 0.56	4.73 ± 0.54	4.75 ± 0.55	0.7484
Hemoglobin, g/L (IQR)	145.00 (136.00–157.00)	146.00 (136.50–157.00)	145.10 (136.00–157.00)	0.4197
Platelet count, *10^{12} (IQR)	255.00 (215.00–312.80)	257.00 (210.00–312.00)	214.50 (175.25–255.00)	0.9621
Lymphocyte count, *10^9 (IQR)	2.02 (1.54–2.51)	1.97 (1.55–2.50)	1.54 (1.19–2.01)	0.5901
Neutrophil count, *10^9 (IQR)	5.34 (4.22–6.78)	5.48 (4.41–7.00)	4.23 (3.33–5.38)	0.3002
Neutrophil to lymphocyte ratio (IQR)	2.63 (1.94–3.76)	2.60 (1.88–4.27)	2.62 (1.93–3.84)	0.5424
Platelet-to-lymphocyte ratio (IQR)	138.24 (103.82–177.23)	135.83 (104.29–183.37)	138.23 (104.01–178.40)	0.8353
BUN, mmol/L (IQR)	5.02 (4.19–5.92)	4.66 (3.97–5.95)	4.96 (4.14–5.92)	0.11
Creatinine, μmol/L (IQR)	65.1 (55.2–75.8)	65.2 (53.6–75.55)	65.1 (55–75.78)	0.5847
Glucose on admission, mmol/L (IQR)	6.57 (5.68–8.00)	6.83 (5.8–9.26)	6.59 (5.69–8.10)	0.1123
CRP, mg/L (IQR)	5.35 (3.02–5.80)	5.32 (3.01–5.98)	5.34 (3.02–5.83)	0.5633
D-dimer level on admission mg/L (IQR)	0.10 (0.06–0.16)	0.10 (0.05–0.16)	0.10 (0.06–0.16)	0.6774
TC mmol/L (IQR)	3.91 (3.27–4.58)	4.01 (3.19–4.70)	3.94 (3.26–4.59)	0.5695
Triglyceride, mmol/L (IQR)	1.34 (1.00–1.80)	1.33 (1.01–1.89)	1.34 (1.00–1.80)	0.7775
HDL-C mmol/L (IQR)	0.96 (0.83–1.10)	0.93 (0.82–1.10)	0.96 (0.82–1.10)	0.4494
LDL-C, mmol/L (IQR)	2.41 (1.89–2.95)	2.56 (1.97–3.10)	2.43 (1.89–2.97)	0.1922

(Continued)

TABLE 1 (Continued)

Variables	Clinically stable ($n = 891$)	END ($n = 129$)	Total ($n = 1,020$)	*P*-value
ApoA1, mmol/L (IQR)	1.12 (1.04–1.18)	1.12 (1.03–1.20)	1.12 (1.04–1.19)	0.0771
Apo B, mmol/L (IQR)	0.86 (0.78–1.04)	0.96 (0.83–1.09)	0.78 (0.67–0.87)	<0.0001
ApoA1 to ApoB ratio, (IQR)	1.57 (1.29–1.66)	1.47 (1.14–1.52)	1.57 (1.27–1.65)	0.0001
Homocysteine, μmol/L (IQR)	19.60 (14.30–32.20)	19.00 (15.25–27.25)	19.5 (14.40–31.80)	0.8667
HbA1c, % (IQR)	5.90 (5.60–6.30)	6.00 (5.60–6.42)	5.90 (5.60–6.30)	0.5895
Clinical features				
Prestroke mRS score, (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	0.879
Time from onset to presentation, h (IQR)	10 (6–18)	9 (4–15)	10 (5–17)	0.096
Admission NIHSS score, (IQR)	3 (1–4)	4 (2–5)	3 (1–4)	<0.0001
Motor arm NIHSS score, (IQR)	1 (0–1)	1 (0–2)	1 (0–1)	<0.0001
Motor leg NIHSS score, (IQR)	1 (0–1)	1 (0–2)	1 (0–1)	<0.0001
Dysarthria NIHSS score, (IQR)	1 (0–1)	1 (0–1)	1 (0–1)	0.1002
Sensory NIHSS score (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	0.9652
Facial palsy NIHSS score, (IQR)	1 (1)	1 (1)	1 (1)	0.001
Time to onset of END, hours (IQR)	–	16 (5–25)	–	–
Treatment of the acute phase, *n* (%)				
Alteplase	66 (7.41)	17 (13.18)	83 (8.14)	0.0251
Loading DAPT	507 (56.90)	67 (51.94)	574 (56.27)	0.288
SAPT	373 (41.86)	51 (39.53)	424 (41.57)	0.616
Lipid-lowering drugs	877 (98.43)	129 (100)	1,006 (98.63)	0.1517
Argatroban	119 (13.36)	8 (6.20)	127 (12.45)	0.0214
Lesion location, *n* (%)				
Basal ganglia	171 (19.19)	43 (33.33)	214 (20.98)	0.0002
Internal capsule	136 (15.26)	31 (24.03)	167 (16.37)	0.0119
Lateral ventricle	208 (23.34)	18 (13.95)	226 (22.16)	0.0164
Centrum semiovale	70 (7.86)	5 (3.88)	75 (7.35)	0.1055
Pons	143 (16.05)	28 (21.71)	171 (16.76)	0.108
Thalamus	163 (18.29)	4 (3.10)	167 (16.37)	<0.0001
Culprit vessel supplying the basal ganglia, *n* (%)				
Lenticulostriate artery	507 (56.90)	91 (70.54)	598 (58.63)	0.0033
Posterior choroidal artery	5 (0.56)	0 (0.00)	5 (0.49)	0.3937
The recurrent artery of Heubner	3 (0.34)	1 (0.78)	4 (0.39)	0.4564
Stroke subtype, *n* (%)				
Lacunar infarction	616 (69.14)	44 (34.11)	660 (64.71)	<0.0001
Branch atheromatous disease	275 (30.86)	85 (65.89)	360 (35.29)	<0.0001
Maximal axial infarct diameter, *n* (%)				<0.0001

(Continued)

TABLE 1 (Continued)

Variables	Clinically stable ($n = 891$)	END ($n = 129$)	Total ($n = 1,020$)	*P*-value
<15 mm	684 (76.76)	61 (47.28)	745 (73.04)	
15–20 mm	144 (16.16)	38 (29.46)	182 (17.84)	
>20 mm	63 (7.07)	30 (23.26)	93 (9.12)	
Layers of cerebral infarction lesions, (IQR)	2 (1–3)	2 (2–3)	2 (1–3)	<0.0001
WMH, (IQR)	2 (1–3)	2 (1–3)	2 (1–3)	0.0513

BMI, body mass index; WBC, white blood cell; RBC, red blood cell; BUN, blood urea nitrogen; CRP, C-reactive protein; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; Apo A1, apolipoproteins A1; ApoB, apolipoproteins B; HbA1c, glycated hemoglobin A 1c; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; DAPT, dual antiplatelet therapy; SAPT, single antiplatelet therapy; WMH, white matter hyperintensities.

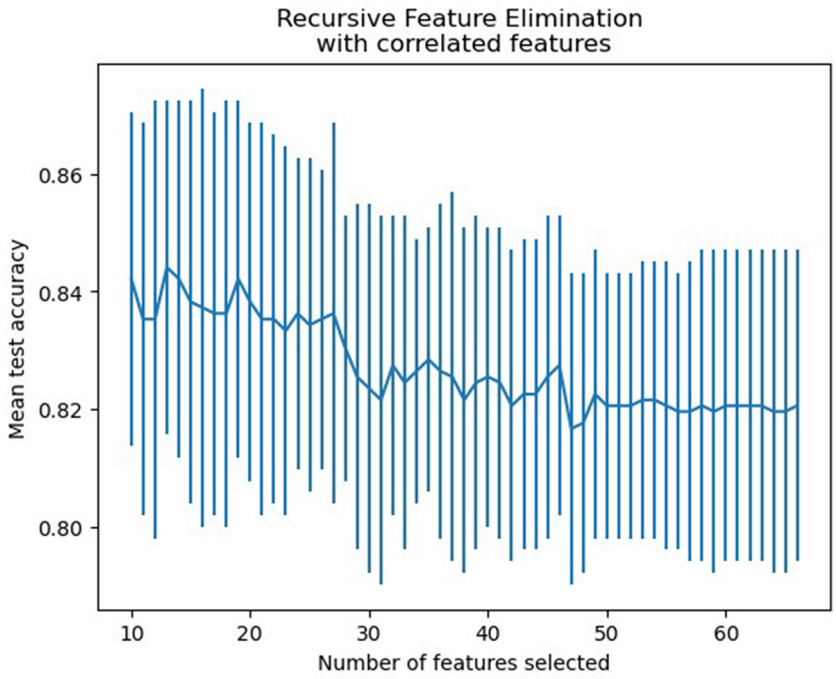


FIGURE 2
Features selection accuracy curve (the accuracy achieved its peak when the number of variables was 13).

arteries, maximum axial infarct diameter (measured at <15 mm), and stroke subtype. Figure 2 shows how accuracy varies with changes in variables.

3.3 Model performance

The training set for model development comprised 714 patients, including 90 with an END outcome, while the testing set for evaluating model performance consisted of 306 patients, 39 of whom experienced an END outcome. Supplementary Table 2 provides a detailed overview of the features selected for both datasets. We employed seven ML algorithms, namely LR, RF, AdaBoost, GBDT, HGB, XGBoost, and CatBoost, to determine the most effective predictive model. Table 2 displays the area under the curve (AUC) for these seven ML algorithms on both the training and testing datasets. Additionally, it offers a comprehensive analysis of accuracy, F1-score, Matthew's correlation coefficient (MCC), specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), and Youden's index for these algorithms on the testing dataset. The GBDT model achieved the highest AUC value at 0.914—an essential measure for evaluating predictive model performance, followed by the CatBoost, XGBoost, HGB, RF, AdaBoost, and LR models (0.8923, 0.8807, 0.876, 0.8639, 0.8184, 0.7838, respectively). Figure 3 illustrates the ROC curve and AUC for each ML classifier in the testing dataset. In conclusion, the GBDT model outperformed the other six ML algorithms, suggesting its superior effectiveness in our study. The confusion matrix for GBDT is shown in Figure 4.

10.3389/fneur.2024.1368902

TABLE 2 Summary of prediction results of six ML algorithms based on the training and testing dataset.

	AUC (95%CI) of the test set	AUC (95% CI) of the training set	Accuracy	*F1*-score	MCC	Specificity	Sensitivity	PPV	NPV	Youden's index
LR	0.7838 (0.6956, 0.8720)	0.7510 (0.6905, 0.8115)	0.8758	0.05	0.1498	1	0.0256	1	0.8754	0.0256
RF	0.8669 (0.7928, 0.9410)	0.9941 (0.9829, 1.005)	0.9118	0.5091	0.5265	0.9965	0.3590	0.8750	0.9138	0.3515
AdaBoost	0.8184 (0.7351, 0.9017)	1 (1, 1)	0.8987	0.4364	0.4385	0.985	0.3077	0.75	0.9069	0.2927
GDBT	0.914 (0.8524, 0.9757)	0.8980 (0.8544, 0.9417)	0.915	0.5	0.5511	1	0.3333	1	0.9113	0.3333
HGB	0.8760 (0.8040, 0.9480)	0.9206 (0.8814, 0.9597)	0.915	0.5	0.5511	1	0.3333	1	0.9113	0.3333
Xgboost	0.8807 (0.8099, 0.9516)	0.9238 (0.8854, 0.9622)	0.915	0.5	0.5511	1	0.3333	1	0.9113	0.3333
CatBoost	0.8923 (0.8243, 0.9602)	0.9550 (0.9248, 0.9851)	0.9085	0.44	0.5052	1	0.282	1	0.905	0.282

3.4 Interpretation of the machine learning model

To evaluate the significance of each feature in the predictive model, we applied the SHAP method to the GBDT model in the testing dataset. The SHAP value analysis revealed that the most impactful features were apoB, TC, BUN, LDL-C, and a maximal axial infarct diameter of <15 mm. Following these features were the stroke subtype, the basal ganglia, the thalamus, argatroban, single antiplatelet therapy, and loading dual antiplatelet therapy. Atrial fibrillation and posterior choroidal arteries also contributed to the prediction model but exhibited lower SHAP values. Figure 5 shows the SHAP summary plot for the GBDT model in the testing dataset, where each dot represents an individual case; the color indicates the feature's value (red for higher, blue for lower). Notably, a higher SHAP value for a feature indicates a greater likelihood of END occurrence. Figure 6 displays the ranking of feature importance based on SHAP values.

4 Discussion

ML techniques that are integral to artificial intelligence have gained substantial attention and are increasingly employed in medical research for tasks such as screening, diagnosis, and prognosis. Recent studies (15, 35) have investigated the use of ML algorithms in predicting END in patients with acute minor stroke and atrial fibrillation-related stroke. Although these studies yielded promising results with ML, they did not specifically focus on predicting END in patients with PAI—a stroke subtype with a high incidence of END. Our study aimed to use seven ML models to predict END in PAI patients. To address the inherent "black box" nature of ML, we employed the SHAP method to elucidate the predictions of the most effective model, ensuring both the model's performance and its clinical interpretability. This approach enabled the effective communication of information through intuitive visual tools, thereby enhancing clinicians' comprehension of the model's decision-making process and aiding in the clinical application of the prediction results.

In our research, we demonstrated that interpretable machine learning techniques can effectively predict END and personalize predictions for individual patients. The results showed that the GBDT model surpassed six other ML algorithms in terms of AUC and accuracy. Furthermore, the five most important variables associated with END prediction were identified as apoB, TC, BUN, LDL-C, and a maximum axial infarct diameter of <15 mm. Previous studies have indicated that the END in single subcortical infarctions, including lacunar stroke, is influenced by various factors such as capsular warning syndrome, higher mean arterial pressure at admission, the location of the infarct in the ventral pons, and the extent of hypoperfusion lesion on perfusion-weighted imaging (17). Other factors, such as the initial NIHSS score, pulsatility index, parent artery disease, and neutrophil-to-lymphocyte ratio, also play significant roles (9). In our study, END in PAI was determined by multiple factors, distinguishing it from previous research. The differences between studies may be attributed to variations in study populations, the

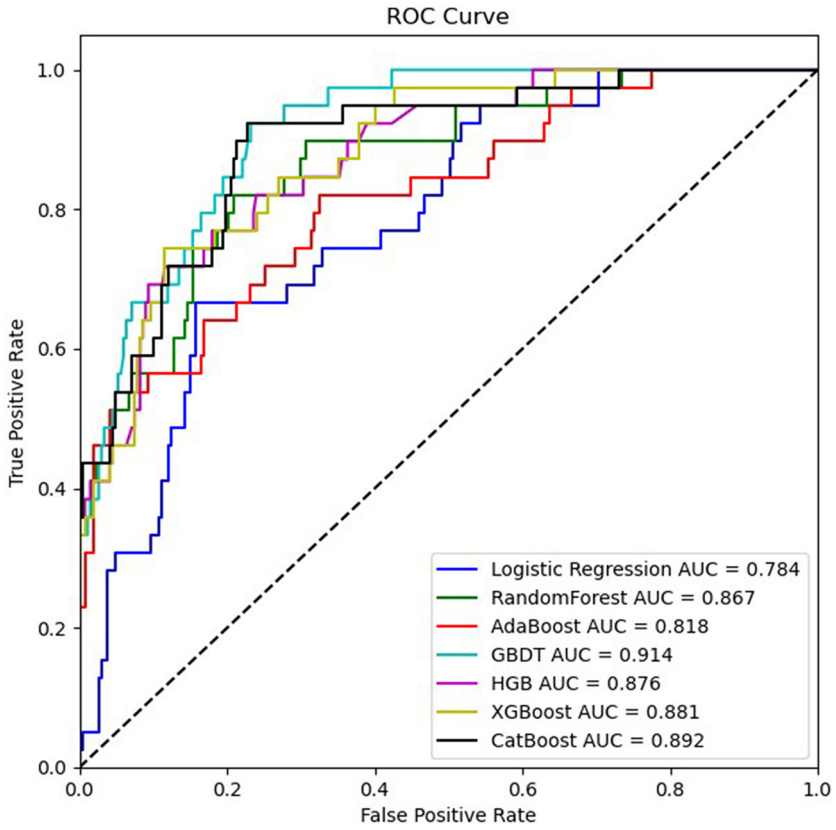


FIGURE 3
ROC curves of seven ML algorithms based on variables in the testing dataset.

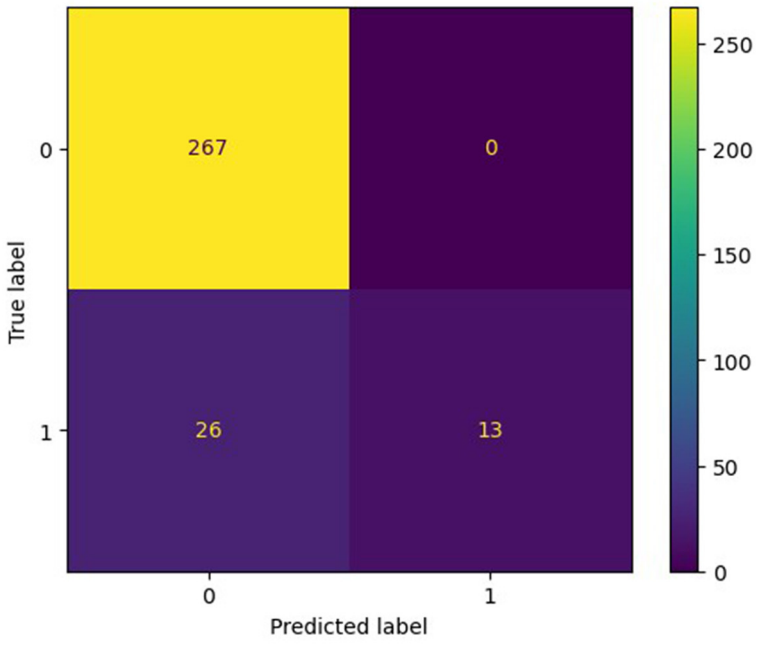


FIGURE 4
The confusion matrix of the most effective model, GBDT.

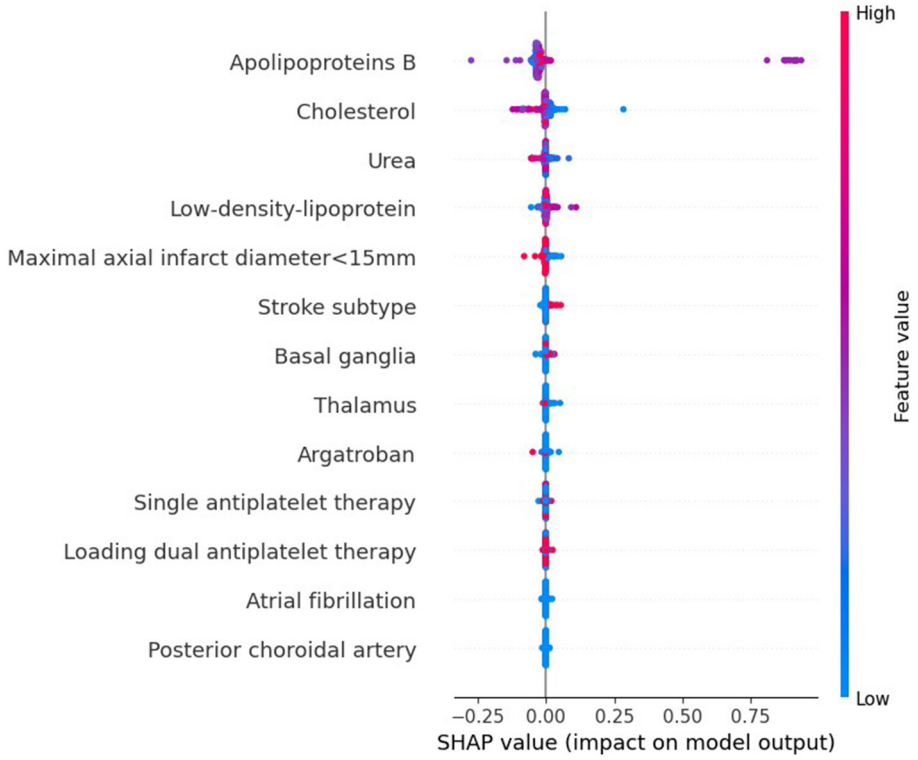


FIGURE 5
The SHAP summary plot on the test data derived from the optimal prediction model, GBDT.

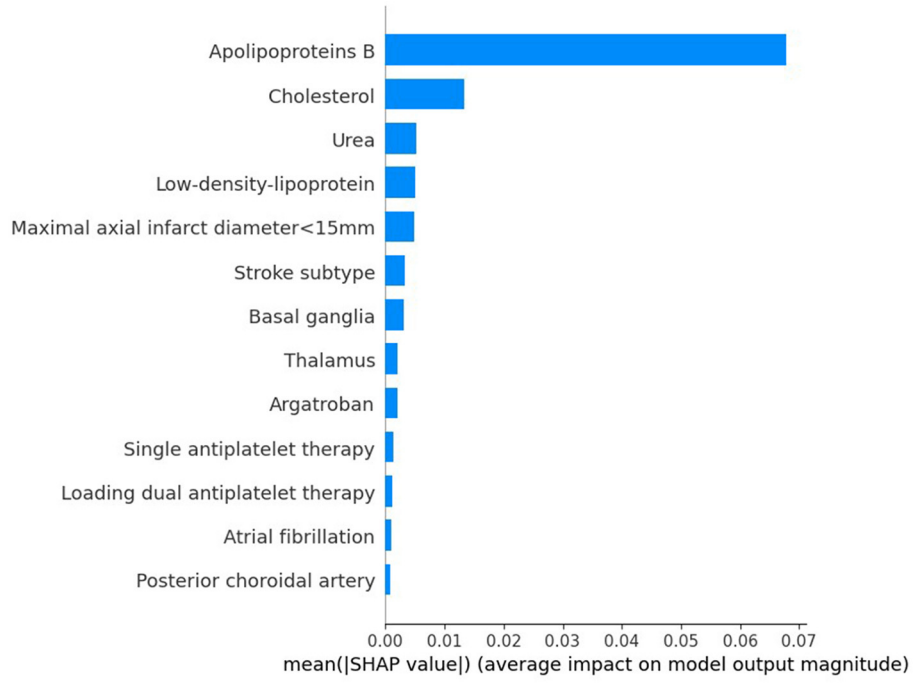


FIGURE 6
Ranking of the features' importance indicated by SHAP analysis of the best prediction model, GBDT.

influence of different statistical methodologies, and the inclusion of diverse variables.

Initially, the END group exhibited dyslipidemia, characterized by higher levels of apoB, TC, and LDL-C compared to the clinically stable group. ApoB, an essential structural component of atherogenic lipoprotein particles such as LDL, lipoprotein, and triglyceride-rich lipoproteins, is acknowledged as a predictor of ischemic cerebrovascular events in patients with preexisting cardiovascular diseases (36). LDL-C, a commonly used clinical lipid marker for assessing lipid-associated risk, including ischemic stroke, has been linked to a reduced frequency of cardiovascular events at lower levels (37). Moreover, previous studies have shown a positive association between TC levels and ischemic stroke risk (38). However, the exact relationship between these lipids and the outcome of END is not yet fully understood. Elevated apoB levels may increase endothelial permeability to LDL, and there is a positive correlation between hypercholesterolemia and apoB generation within plaques. High apoB levels facilitate the penetration of particles into the arterial wall, leading to localized accumulation within the subendothelium. This process heightens the susceptibility to modifications in the artery wall, contributing to the onset of atherosclerosis and plaque progression. Thus, this study suggests that high levels of apoB, TC, and LDL-C, particularly apoB, are linked to END in PAI patients. Previous studies have revealed that a high baseline level of apoB correlates with an increased risk of major adverse cardiovascular events in acute coronary syndrome (39). This study observed lower BUN levels in the END group, with reduced BUN levels correlating with an increased risk of developing END, aligning with previous research (40), which suggested that lower BUN levels were observed in patients with progressive infarction in the anterior circulation and small subcortical infarction.

Previous studies have reported that END is more frequently observed in patients with BAD (41), a trend also evident in our study, suggesting an association between BAD-related PAI and the development of END. We also found that the proportion of maximum axial infarct diameters of <15 mm was lower in the END group than in the clinically stable group. This finding implies that smaller infarct diameters (<15 mm), typically linked to LI, are negatively correlated with the occurrence of END. The infarction location significantly influences functional limitations, clinical progression, and patient outcomes (42). Previous studies have indicated that infarct locations within the brainstem, corona radiata, and lenticulostriate artery area, including the internal capsule, are associated with an increased risk of END (4, 43–45). In our study, lesions in the basal ganglia (excluding the internal capsule) and thalamus were identified as predictors of END, with the former located in the lenticulostriate artery area, aligning with previous studies, and the latter showing comparatively favorable outcomes in terms of mortality and permanent motor deficits.

In the management of acute ischemic stroke, antiplatelet therapy is a fundamental therapeutic strategy. Several studies have indicated that DAPT reduces the risk of END (17, 46, 47). In our study, the utilization of antiplatelet therapy in the END group was less frequent than in the clinically stable group. However, our findings suggest that DAPT does not show a clear advantage over SAPT in preventing END. Moreover, we observed that combining argatroban with antiplatelet therapy was associated with a decreased risk of END in patients with PAI, consistent with previous research (48). The contribution of the AF and posterior choroidal artery to the model was found to be minimal.

Our analysis demonstrated that ML models, particularly the GBDT algorithm, showed promising outcomes in predicting END in PAI patients. GBDT, a sophisticated ML algorithm, integrates multiple decision trees to develop a more accurate and robust model. This algorithm effectively handles both continuous and categorical variables, shows a lower susceptibility to overfitting compared to more complex models, and adeptly manages missing data. Considering the inherent multivariate heterogeneity and noise in clinical research data, such as demographic information, laboratory findings, and radiological results available upon hospital arrival, it is critical to choose variables based on prior knowledge. Consequently, we used recursive feature elimination with cross-validation (RFECV) to select 13 predictors for the model, enhancing prediction accuracy.

A strength of this study is that the model performance and clinical interpretability were ensured by using the SHAP algorithm, which was effectively presented to users through easy-to-use visualization tools. Clinicians could better understand the model's decision-making process, thus facilitating the clinical application of prediction results. In addition, our ML model for interpreting predictions was based on a large number of variables, including demographics and laboratory/radiological data obtained from real-world clinical situations at hospital admission. Finally, we demonstrated the potential of interpretable machine learning methods for predicting END in PAI patients and personalizing these predictions within patient populations.

The limitations of our study are as follows: First, the data used in this study is retrospective in nature and sourced from a single center. The retrospective nature of the data may have introduced recall and selection biases to varying degrees. Therefore, it is necessary to use more datasets and conduct prospective multicenter clinical trials to further verify the results and enhance the model's accuracy. Second, we only performed internal validation for dataset validation, but external validation is needed to assess the robustness of the ML model further. Third, missing values were handled using mean interpolation, which inevitably introduces a degree of bias. However, if missing values are removed, some selection bias cannot be completely avoided. Fourth, the exclusion of omics data from the study may potentially limit the predictive performance to some extent.

5 Conclusion

We have demonstrated that seven ML models, particularly the GBDT model, can accurately predict END in PAI patients. However, further research with a larger cohort is essential to validating the model's accuracy. Additionally, the predictive efficacy of this model merits exploration in prospective clinical studies.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Jincheng People's Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

WL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LJ: Writing – review & editing. LX: Writing – review & editing. FY: Writing – review & editing. ZG: Writing – review & editing. JL: Writing – review & editing. DZ: Writing – review & editing, Data curation. YL: Data curation, Writing – review & editing. HX: Data curation, Writing – review & editing. HC: Writing – review & editing, Formal analysis. JH: Writing – review & editing. SL: Writing – review & editing. HL: Formal analysis, Writing – review & editing.

Funding

The author (s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1368902/full#supplementary-material

References

1. Liao X, Feng S, Wang Y, Pan Y, Chen W, Qu H, et al. Tirofiban combined with aspirin in the treatment of acute penetrating artery territory infarction (STRATEGY): protocol for a multicentre, randomised controlled trial. *Stroke Vasc Neurol.* (2023) 9:75–81. doi: 10.1136/svn-2022-002284

2. Yaghi S, Raz E, Yang D, Cutting S, Mac Grory B, Elkind MS, et al. Lacunar stroke: mechanisms and therapeutic implications. *J Neurol Neurosurg Psychiatry.* (2021) jnnp-2021-326308. doi: 10.1136/jnnp-2021-326308

3. Nah H-W, Kang D-W, Kwon SU, Kim JS. Diversity of single small subcortical infarctions according to infarct location and parent artery disease: analysis of indicators for small vessel disease and atherosclerosis. *Stroke.* (2010) 41:2822–7. doi: 10.1161/STROKEAHA.110.599464

4. Jin D, Yang J, Zhu H, Wu Y, Liu H, Wang Q, et al. Risk factors for early neurologic deterioration in single small subcortical infarction without carrier artery stenosis: predictors at the early stage. *BMC Neurol.* (2023) 23:83. doi: 10.1186/s12883-023-03128-3

5. Caplan LR. Lacunar infarction and small vessel disease: pathology and pathophysiology. *J Stroke.* (2015) 17:2–6. doi: 10.5853/jos.2015.17.1.2

6. Regenhardt RW, Das AS, Lo EH, Caplan LR. Advances in understanding the pathophysiology of lacunar stroke: a review. *JAMA Neurol.* (2018) 75:1273–81. doi: 10.1001/jamaneurol.2018.1073

7. Jeong H-G, Kim BJ, Yang MH, Han M-K, Bae H-J. Neuroimaging markers for early neurologic deterioration in single small subcortical infarction. *Stroke.* (2015) 46:687–91. doi: 10.1161/STROKEAHA.114.007466

8. Jiang J, Huang X, Zhang Y, Deng W, Shen F, Liu J. Total MRI burden of cerebral vessel disease correlates with the progression in patients with acute single small subcortical strokes. *Brain Behav.* (2019) 9:e01173. doi: 10.1002/brb3.1173

9. Nam K-W, Kwon H-M, Lee Y-S. Different predictive factors for early neurological deterioration based on the location of single subcortical infarction: early prognosis in single subcortical infarction. *Stroke.* (2021) 52:3191–8. doi: 10.1161/STROKEAHA.120.032966

10. Han X-J. Relationship between lesion patterns of single small infarct and early neurological deterioration in the perforating territory. *Eur Rev Med Pharmacol Sci.* (2017) 21:3642–8.

11. Helleberg BH, Ellekjaer H, Indredavik B. Outcomes after early neurological deterioration and transitory deterioration in acute ischemic stroke patients. *Cerebrovasc Dis.* (2016) 42:378–86. doi: 10.1159/000447130

12. Chen SD, You J, Yang XM, Gu HQ, Huang XY, Liu H, et al. Machine learning is an effective method to predict the 90-day prognosis of patients with transient ischemic attack and minor stroke. *BMC Med Res Methodol.* (2022) 22:195. doi: 10.1186/s12874-022-01672-z

13. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* (2019) 380:1347–58. doi: 10.1056/NEJMra1814259

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) 50:1263–5. doi: 10.1161/STROKEAHA.118.024293

15. Kim S-H, Jeon E-T, Yu S, Oh K, Kim CK, Song T-J, et al. Interpretable machine learning for early neurological deterioration prediction in atrial fibrillation-related stroke. *Sci Rep.* (2021) 11:20610. doi: 10.1038/s41598-021-99920-7

16. Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* (2013) 12:822–38. doi: 10.1016/S1474-4422(13)70124-8

17. Vynckier J, Maamari B, Grunder L, Goeldlin MB, Meinel TR, Kaesmacher J, et al. Early neurologic deterioration in lacunar stroke: clinical and imaging predictors and association with long-term outcome. *Neurology.* (2021) 97:e143746. doi: 10.1212/WNL.0000000000012661

18. Petrone L, Nannoni S, Del Bene A, Palumbo V, Inzitari D. Branch atheromatous disease: a clinically meaningful, yet unproven concept. *Cerebrovasc Dis.* (2016) 41:87–95. doi: 10.1159/000442577

19. van Leijsen EMC, van Uden IWM, Ghafoorian M, Bergkamp MI, Lohner V, Kooijmans ECM, et al. Nonlinear temporal dynamics of cerebral small vessel disease. *Neurology.* (2017) 89:1569–77. doi: 10.1212/WNL.0000000000004490

20. Nusinovici S, Tham YC, Chak Yan MY, Wei Ting DS Li J, Sabanayagam C, Wong TY, et al. Logistic regression was as good as machine learning for predicting major chronic diseases. *J Clin Epidemiol.* (2020) 122:56–69. doi: 10.1016/j.jclinepi.2020.03.002

21. Yang L, Wu H, Jin X, Zheng P, Hu S, Xu X, et al. Study of cardiovascular disease prediction model based on random forest in eastern China. *Sci Rep.* (2020) 10:5245. doi: 10.1038/s41598-020-62133-5

22. Zheng Z, Yang Y. Adaptive boosting for domain adaptation: toward robust predictions in scene segmentation. *IEEE Trans Image Process.* (2022) 31:5371–82. doi: 10.1109/TIP.2022.3195642

23. Seto H, Oyama A, Kitora S, Toki H, Yamamoto R, Kotoku J, et al. Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data. *Sci Rep.* (2022) 12:15889. doi: 10.1038/s41598-022-20149-z

24. Velichko A, Huyut MT, Belyaev M, Izotov Y, Korzun D. Machine learning sensors for diagnosis of COVID-19 disease using routine blood values for internet of things application. *Sensors.*(2022) 22:7886. doi: 10.3390/s22207886

25. Nhat-Duc H, Van-Duc T. *Comparison of Histogram-Based Gradient Boosting Classification Machine, Random Forest, and Deep Convolutional Neural Network for Pavement Raveling Severity Classification.* Automation in Construction (2023). Available online at: https://www.nstl.gov.cn/paper_detail.html?id=4d53926ab8e76e705f898065f8a78833 (accessed April 19, 2024).

26. Chen T, Guestrin C. XGBoost: a scalable tree boosting system: In: *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM (2016). doi: 10.1145/2939672.2939785

27. Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. *arXiv preprint* arXiv:1810.11363 (2018). doi: 10.48550/arXiv.1810.11363

28. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. *J Big Data.* (2020) 7:94. doi: 10.1186/s40537-020-00369-8

29. Tran BX, Ha GH, Nguyen LH, Vu GT, Hoang MT, Le HT, et al. Studies of novel coronavirus disease 19 (COVID-19) pandemic: a global analysis of literature. *Int J Environ Res Public Health.* (2020) 17:4095. doi: 10.3390/ijerph17114095

30. Wang R, Zhang J, Shan B, He M, Xu J. XGBoost machine learning algorithm for prediction of outcome in aneurysmal subarachnoid hemorrhage. *Neuropsychiatr Dis Treat.* (2022) 18:659–67. doi: 10.2147/NDT.S349956

31. Hou N, Li M, He L, Xie B, Wang L, Zhang R, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med.* (2020) 18:462. doi: 10.1186/s12967-020-02620-5

32. Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. *Comput Biol Med.* (2019) 112:103375. doi: 10.1016/j.compbiomed.2019.103375

33. Yu L, Zhao Y, Wang H, Sun T-L, Murphy TE, Tsui K-L. Assessing elderly's functional balance and mobility via analyzing data from waist-mounted tri-axial wearable accelerometers in timed up and go tests. *BMC Med Inform Decis Mak.* (2021) 21:108. doi: 10.1186/s12911-021-01463-4

34. Lundberg S, Lee SI. *A Unified Approach to Interpreting Model Predictions.* Nips (2017). Available online at: http://www.xueshufan.com/publication/2618851150 (accessed November 6, 2023).

35. Sung SM, Kang YJ, Cho HJ, Kim NR, Lee SM, Choi BK, et al. Prediction of early neurological deterioration in acute minor ischemic stroke by machine learning algorithms. *Clin Neurol Neurosurg.* (2020) 195:105892. doi: 10.1016/j.clineuro.2020.105892

36. Koren-Morag N, Goldbourt U, Graff E, Tanne D. Apolipoproteins B and AI and the risk of ischemic cerebrovascular events in patients with pre-existing atherothrombotic disease. *J Neurol Sci.* (2008) 270:82–7. doi: 10.1016/j.jns.2008.02.005

37. Ntaios G, Milionis H. Low-density lipoprotein cholesterol lowering for the prevention of cardiovascular outcomes in patients with ischemic stroke. *Int J Stroke.* (2019) 14:476–82. doi: 10.1177/1747493019851283

38. Zhang Y, Tuomilehto J, Jousilahti P, Wang Y, Antikainen R, Hu G. Total and high-density lipoprotein cholesterol and stroke risk. *Stroke.* (2012) 43:646778. doi: 10.1161/STROKEAHA.111.646778

39. Hagström E, Steg PG, Szarek M, Bhatt DL, Bittner VA, Danchin N, et al. Apolipoprotein B, residual cardiovascular risk after acute coronary syndrome, and effects of alirocumab. *Circulation.* (2022) 146:657–72. doi: 10.1161/CIRCULATIONAHA.121.057807

40. Lin J, Mao X, Liao Y, Luo S, Huang Q, Song Z, et al. A lesion extending three or more slices as a predictor of progressive infarction in anterior circulation small subcortical infarction. *Front Neurol.* (2022) 13:926187. doi: 10.3389/fneur.2022.926187

41. Sun S, Wang Y, Wang Y, Men X, Bao J, Hu X, et al. Lipid and hyperglycemia factors in first-ever penetrating artery infarction, a comparison between different subtypes. *Brain Behav.* (2017) 7:e00694. doi: 10.1002/brb3.694

42. Yu W, Yang J, Liu L, Song W, Zhang Z, Xu M, et al. The value of diffusion weighted imaging in predicting the clinical progression of perforator artery cerebral infarction. *Neuroimage Clin.* (2022) 35:103117. doi: 10.1016/j.nicl.2022.103117

43. Liu H, Liu K, Zhang K, Zong C, Yang H, Li Y, et al. Early neurological deterioration in patients with acute ischemic stroke: a prospective multicenter cohort study. *Ther Adv Neurol Disord.* (2023) 16:17562864221147743. doi: 10.1177/17562864221147743

44. Li H, Dai Y, Wu H, Luo L, Wei L, Zhou L, et al. Predictors of early neurologic deterioration in acute pontine infarction. *Stroke.* (2020) 51:637–40. doi: 10.1161/STROKEAHA.119.027239

45. Berberich A, Schneider C, Herweh C, Hielscher T, Reiff T, Bendszus M, et al. Risk factors associated with progressive lacunar strokes and benefit from dual antiplatelet therapy. *Eur J Neurol.* (2020) 27:817–24. doi: 10.1111/ene.14159

46. He F, Xia C, Zhang J-H, Li X-Q, Zhou Z-H, Li F-P, et al. Clopidogrel plus aspirin versus aspirin alone for preventing early neurological deterioration in patients with acute ischemic stroke. *J Clin Neurosci.* (2015) 22:83–6. doi: 10.1016/j.jocn.2014.05.038

47. Yi X, Zhou Q, Wang C, Lin J, Chai Z. Aspirin plus clopidogrel may reduce the risk of early neurologic deterioration in ischemic stroke patients carrying CYP2C19*2 reduced-function alleles. *J Neurol.* (2018) 265:2396–403. doi: 10.1007/s00415-018-8998-1

48. Xu J, Xu X, Wang H, He L, Liu Q, Du Y, et al. Dual antiplatelet therapy plus argatroban prevents early neurological deterioration in branch atherosclerosis disease. *Stroke.* (2022) 53:e19–20. doi: 10.1161/STROKEAHA.121.036356

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 30 May 2024
DOI 10.3389/fneur.2024.1414959

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Chun-Jen Lin,
Taipei Veterans General Hospital, Taiwan
Zhongming Qiu,
Xinqiao Hospital, China

*CORRESPONDENCE
Shudong Liu
✉ shudongliu@live.cn

†These authors have contributed equally to this work

‡These authors share last authorship

RECEIVED 09 April 2024
ACCEPTED 17 May 2024
PUBLISHED 30 May 2024

CITATION
Zhang L, Xue H, Bu X, Liao J, Tang G, Chen Y, Zhao L, Yang D, Liu L and Liu S (2024) Patchy profile sign in RAPID software: a specific marker for intracranial atherosclerotic stenosis in acute ischemic stroke. *Front. Neurol.* 15:1414959. doi: 10.3389/fneur.2024.1414959

COPYRIGHT
© 2024 Zhang, Xue, Bu, Liao, Tang, Chen, Zhao, Yang, Liu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Patchy profile sign in RAPID software: a specific marker for intracranial atherosclerotic stenosis in acute ischemic stroke

Lingwen Zhang[1,2†], Hua Xue[1,2†], Xiaoqing Bu[3†], Juan Liao[2†], Ge Tang[1,2], Yu Chen[1,2], Libo Zhao[1,2], Deyu Yang[1,2], Li Liu[2,4‡] and Shudong Liu[1,2*‡]

[1]Department of Neurology, Yongchuan Hospital of Chongqing Medical University, Chongqing, China, [2]Chongqing Key Laboratory of Cerebrovascular Disease Research, Yongchuan Hospital of Chongqing Medical University, Chongqing, China, [3]Department of Epidemiology, School of Public Health, Chongqing Medical University, Chongqing, China, [4]Department of Health Management, Yongchuan Hospital of Chongqing Medical University, Chongqing, China

Purpose: Identifying the etiology of acute ischemic stroke (AIS) before endovascular treatment (EVT) is important but challenging. In CT perfusion imaging processed by perfusion software, we observed a phenomenon called patchy profile sign (PPS), that is, the hypoperfusion morphology in RAPID software is a discontinuous sheet pattern. This phenomenon is predominantly observed in patients diagnosed with intracranial atherosclerotic stenosis (ICAS). The study intends to assess whether the PPS can be used to differentiate ICAS from intracranial embolism.

Method: Patients with AIS due to M1 segment occlusion of the MCA who underwent mechanical thrombectomy were retrospectively enrolled. The receiver operating characteristic (ROC) curve analysis was performed to assess the value of PPS in predicting ICAS. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the PPS for prediction of ICAS were calculated.

Results: A total of 51 patients were included in the study. The PPS was observed in 10 of 19 (52.6%) patients with ICAS, and in 2 of 32 (6.3%) patients with intracranial embolism ($p < 0.001$). Interobserver agreement for identifying PPS was excellent ($\kappa = 0.944$). The sensitivity, specificity, PPV, NPV, and accuracy of the PPS for predicting ICAS were 52.6, 93.8, 83.3, 76.9, and 78.4%, respectively.

Conclusion: The PPS on RAPID software is an imaging marker with high specificity for ICAS. Larger sample sizes are imperative to validate the findings.

KEYWORDS

ischemic stroke, intracranial embolism, perfusion, software, atherosclerotic, computed tomography

1 Introduction

Acute ischemic stroke (AIS) resulting from large vessel occlusion (LVO) stands as a prominent cause of global morbidity and mortality. Mechanical thrombectomy (MT) has emerged as a pivotal intervention, notably improving outcomes in LVO (1–6). Nevertheless, the prognosis of certain patients undergoing EVT for LVO remains suboptimal, primarily attributable to the

interval between onset and intervention, surgical procedure duration, and the volume of cerebral ischemia during interventional procedures (7). Intracranial atherosclerotic stenosis (ICAS) and embolic etiologies constitute the primary pathogenesis of LVO, with ICAS prevalent in Asian populations (8, 9). Additionally, ICAS-related LVO presents distinct challenges, including lower recanalization rates and extended procedural durations (10–12). Owing to subsequent platelet aggregation, patients with ICAS frequently encounter residual stenosis and reocclusion during mechanical thrombectomy. In cases of reocclusion, considering rescue treatment, such as balloon and stent angioplasty, is advisable (13–16). Clear delineation of pathogenesis before intervention will aid in formulating a personalized treatment strategy and enhance the procedural workflow. Preoperative differentiation between ICAS and embolic LVO remains imperative, yet effective imaging biomarkers are lacking, warranting further investigation.

Perfusion imaging utilizing artificial intelligence (AI) software has become a primary modality for preoperative MT assessment. This modality automatically performs image postprocessing of CTP imaging system-derived images, accurately identifying and quantifying the infarct core and ischemic penumbra (17, 18). Multiple clinical studies have confirmed that patients identified using the RAPID software based on perfusion imaging may derive benefits from endovascular therapy within an extended time window (19, 20). Parameter of perfusion imaging, such as the hypoperfusion intensity ratio (HIR) is valuable in discerning collateral flow in patients with anterior LVO (21). Both the HIR and cerebral blood volume (CBV) index are associated with underlying ICAS and may function as predictors of ICAS before EVT (22).

In our clinical practice, we have documented the patchy profile sign (PPS), an observed phenomenon in some patients where the hypoperfusion morphology manifests as a non-continuous sheet pattern. Specifically, we have observed that PPS is more prone to manifest in patients with ICAS-associated LVO. We postulated that PPS could serve as a valuable imaging marker for predicting ICAS before EVT. This study aims to objectively ascertain whether the PPS on RAPID software can effectively differentiate between ICAS and intracranial embolism before EVT.

2 Materials and methods

The study was approved by the Ethics Committee, and the need for informed consent was waived for the retrospective nature of the study. The procedures of this study adhere to the declaration of Helsinki.

2.1 Study participants

The study retrospectively involved 51 patients selected from our database of consecutive AIS patients who underwent emergency EVT at the Comprehensive Stroke Center from December 2018 to December 2022. Figure 1 depicts the patient screening process. Inclusion criteria comprised individuals aged over 18, admitted within 24 h from symptom onset, undergoing CTP within 24 h of onset, experiencing ischemic stroke due to MCA M1 occlusion, and receiving EVT with successful recanalization (defined as an mTICI grade of 2b-3 or eTICI grade of 2b-3). Exclusion criteria involved individuals with MCA M1 lesions lacking embolism or stenosis, tandem lesions of the MCA, a history of stroke disease, incomplete clinical or imaging data, or poor image quality (concurrent intracranial structural lesions or strenuous activity during image refinement). Collected clinical data of study participants encompassed demographic characteristics (age and gender), stroke risk factors (history of hypertension, diabetes, atrial fibrillation, hyperlipidemia, smoking, and drinking), and clinical characteristics (Intravenous thrombolysis, systolic and diastolic blood pressure, admission random intravenous blood glucose, admission NIHSS, 24-h post-operative NIHSS, admission Glasgow Coma Scale [GCS], ASPECTS, onset to imaging time, onset to puncture time). Data supporting the study can be obtained from the corresponding authors upon reasonable request.

2.2 Imaging data

All stroke patients underwent a comprehensive CT scan, consisting of a non-contrast CT, CT angiography of the head and neck, and CT perfusion, using a 256-slice multi-detector CT scanner (Brilliance iCT). Initially, a non-contrast CT scan of the head was performed to rule out intracranial bleeding, followed by a CT angiogram of the head and neck, and subsequent CT perfusion imaging. Whole-brain helical NCCT (120 kVp,100–350 auto-mAs) was performed with 5-mm section thickness. CT perfusion parameters were obtained in a periodic spiral pattern. A high-pressure syringe was utilized to inject 70–90 mL of the contrast agent iopamidol at a flow rate ranging from 4.0 to 6.0 mL/s. Subsequently, the tube was flushed with 30 mL of physiological saline, and the scan commenced with a 5-s delay. The imaging spanned from the foramen magnum to the level above the lateral ventricle, utilizing an 80 mm collimation, tube voltage of 80 kV, and tube current of 100 mA. The perfusion maps and their associated parameters underwent automated analysis using the RAPID software (iSchemaView, Menlo Park, CA; version 5.0.4). The ischemic core was defined as a tissue volume with cerebral blood flow of <30% on CTP imaging. Hypoperfusion was defined as a volume of tissue of Tmax >6 s on CTP. The mismatch ratio was calculated by dividing the ischemic core volume by the lesion volume with Tmax >6 s. The mismatch volume was calculated by subtracting the ischemic core volume from the lesion volume with a Tmax >6 s. HIR was defined as the ratio of the volume of the "Tmax >10 s" lesion divided by the volume of the "Tmax >6 s" lesion. The CBV index was defined as the ratio of the mean CBV within the "Tmax >6 s" lesion in the ipsilateral hemisphere over the mean CBV of the unaffected brain area. We defined HIR $\leq$ 0.22 and CBV $\geq$ 0.90 as favorable predictors of atherosclerosis based on previous research (22).

Collateral status on CTA was assessed by a straightforward method, which assesses the backfilling of the soft meningeal arteries in the entire MCA ischemic area compared to the contralateral side, defined as (0, minimal; 1, less than 50%; 2, greater than 50%; 3, filling 100% of the ischemic area) (23).

Abbreviations: AIS, acute ischemic stroke; EVT, endovascular treatment; ICAS, intracranial atherosclerotic stenosis; ROC, receiver operating characteristic; PPV, positive predictive value; NPV, negative predictive value; LVO, large vessel occlusion; ACA, anterior cerebral artery; GCS, Glasgow Coma Scale; HIR, hypoperfusion intensity ratio; CBV, cerebral blood volume; PPS, patchy profile sign; AI, artificial intelligence; HMCAS, hyperdense middle cerebral artery sign.

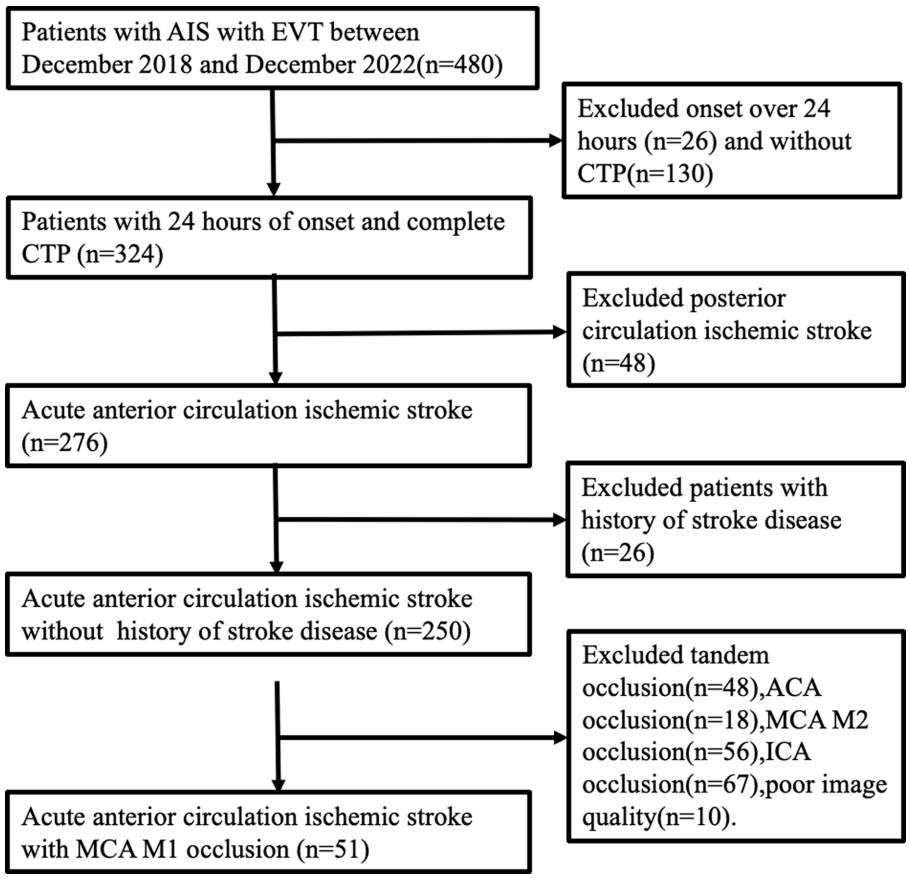


FIGURE 1
Patient selection process. AIS, acute ischemic stroke; EVT, endovascular treatment; ACA, anterior cerebral artery.

2.3 Operational definitions of ICAS and intracranial embolism

ICAS was differentiated from embolism based on the outcome of angiographic findings and endovascular treatment. ICAS was defined as (1) fixed stenosis ≥70% with either angiographically evident impaired perfusion or evidence of reocclusion following sufficient treatment with a stent retriever, and (2) percutaneous transluminal angioplasty and dual antiplatelet therapy were required to maintain effective recanalization (Figures 2C,D). Iatrogenic dissection or vessel wall injury resulting in stenosis was not classified as an ICAS and was excluded from the study. Embolism was classified as there was no or certain focal stenosis and no tendency for reocclusion after clot retrieval. If CTA within 1 week of surgery confirms complete recanalization of the responsible vessel, it is also considered an embolism (22, 24) (Figures 2A,B).

2.4 Definition of the PPS

The positive PPS was defined as (1) a hypoperfusion region comprising two or more comparable scattered patches or primarily contiguous regions with highly irregular edges that do not conform to any geometric or morphological definition. (Figures 3A,B) (2); the aforementioned images accounted for half or more of all images in the presence of hypoperfusion. Small patches of hypoperfusion images considered individually insignificant were not included in the preceding definition. Similarly, images displaying regular edges on one side were also excluded. (Figures 3C,D).

Both physicians were trained on irregular profile and simulant images by means of photographs, and all perfusion images were concealed from any associated information, independently evaluated by two neurologists each with over 3 years of experience. The two neurologists discussed the controversial findings until a consensus was reached. The imaging features of hypoperfusion in RAPID software with and without PPS are illustrated in Figures 4, 5.

2.5 Statistical analysis

Normally distributed data were expressed as mean ± standard deviation (SD), and the student *t* test was used for comparisons between groups. Non-normally distributed data were presented as median (M) with upper and lower quartiles (P25, P75), and the Mann–Whitney *U* test performed the comparison between groups. Categorical variables were expressed as frequencies (percentages, %). The χ^2 tests and Fisher exact tests were used to analyze categorical variables as appropriate. The consistency of the observer in identifying PPS was examined by the Kappa consistency test, with consistency defined as $\kappa = 0.01$ to 0.20, 0.21 to 0.4, 0.41 to 0.6, 0.61

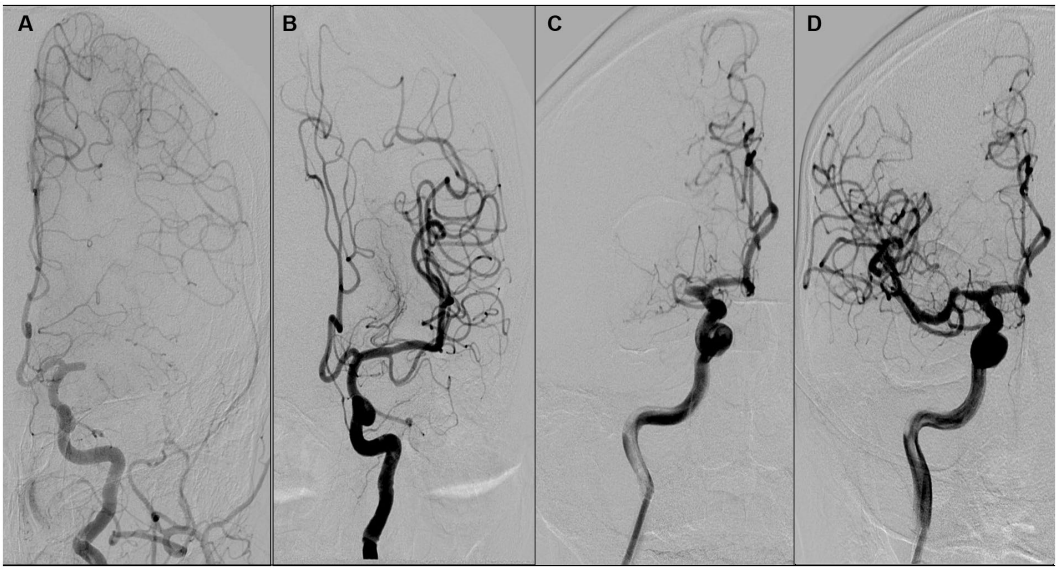


FIGURE 2
Illustration of the definition of atherosclerosis and embolism. **(A,B)** suggest that the etiology of occlusion is an intracranial embolism, while **(C,D)** suggest that occlusion is due to atherosclerosis. A female patient aged 68 years with occlusion of the left middle cerebral artery M1 **(A)**, which was revascularized after arterial suction thrombectomy **(B)**. A male patient aged 68 years with occlusion of the right middle cerebral artery M1 **(C)**, which was revascularized with antiplatelet therapy and stent thrombectomy **(D)**.

to 0.8 and 0.81 to 0.99 indicating slight, fair, moderate, substantial, and excellent interobserver agreement, respectively. The area under the ROC curve (AUC) was performed to assess the value of HIR, CBV, and PPS in predicting ICAS. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated to investigate the diagnostic value of PPS for ICAS. All statistical analyses were performed with IBM SPSS Statistics 26.0 (IBM Corporation, Armonk), and $p < 0.05$ was statistically significant.

3 Results

3.1 Patient characteristics

A total of 480 patients with AIS receiving endovascular treatment were screened, of which 51 met the criteria for inclusion in the analysis. The cause of vascular occlusion was intracranial embolism in 32 patients and ICAS in 19 patients. The PPS was positive in 12 patients and negative in 39 patients.

3.2 Patients with and without PPS

The PPS was observed in 12 (23.5%) of the 51 study patients. The consistency of the two observers in identifying PPS was excellent (Kappa test, $\kappa = 0.944$). Clinical characteristics and imaging features of patients with and without PPS are displayed in Table 1. Compared to patients without PPS, patients with PPS had smaller ischemic core (0 vs. 15.0, $p = 0.001$), smaller hypoperfusion areas (61.5 vs. 158.8, $p < 0.001$), lower HIR index (0.1 vs. 0.5, $p < 0.001$), smaller mismatch volumes (60.6 vs. 128.4, $p < 0.001$), higher CBV index (0.8 vs. 0.7, $p = 0.01$), lower random venous glucose (5.5 vs. 6.9, $p = 0.046$), and less often hyperlipidemia (0 vs. 35.9%, $p = 0.02$). There were no significant differences in gender, age, smoking, drinking, hypertension, diabetes, atrial fibrillation, systolic blood pressure, diastolic blood pressure, intravenous thrombolysis, admission NIHSS, admission GCS, 24-h post-operative NIHSS, onset to imaging time, onset to puncture time, mismatch ratio, fasting blood glucose, CTA collateral score, and ASPECTS between the two groups (all p values >0.05).

3.3 Patients with ICAS vs. patients with intracranial embolism

Patients with ICAS were diagnosed in 19 (37.25%) of the 51 patients. The clinical characteristics and imaging features of patients with ICAS and intracranial embolism are shown in Table 2. Patients with ICAS were more likely to be smoking (52.6% vs. 18.8%, $p = 0.01$), more often than males (78.9% vs. 40.6%, $p = 0.008$), less likely to have atrial fibrillation (10.5% vs. 62.5%, $p < 0.001$) and more often to have positive PPS (52.6% vs. 6.3%, $p < 0.001$). The volume of the ischemic core (0 vs. 15.0, $p = 0.002$), the volume of hypoperfusion (105.3 vs. 154.1, $p = 0.006$), and the HIR index (0.3 vs. 0.5, $p = 0.001$), the admission NIHSS score (12.0 vs. 17.0, $p = 0.003$) were lower in the ICAS group than those in the intracranial embolism group. The CBV index (0.8 vs. 0.7, $p = 0.02$) and ASPECTS (9.0 vs. 7.0, $p = 0.005$) were higher in the ICAS group than in the intracranial embolism group. There were no significant differences in age, drinking, hypertension, diabetes, hyperlipidemia, random venous glucose, intravenous thrombolysis, systolic and diastolic blood pressure, 24-h post-operative NIHSS, admission GCS, onset to imaging time, onset to puncture time, mismatch ratio, mismatch volume, CTA collateral score between the two groups (all p values >0.05).

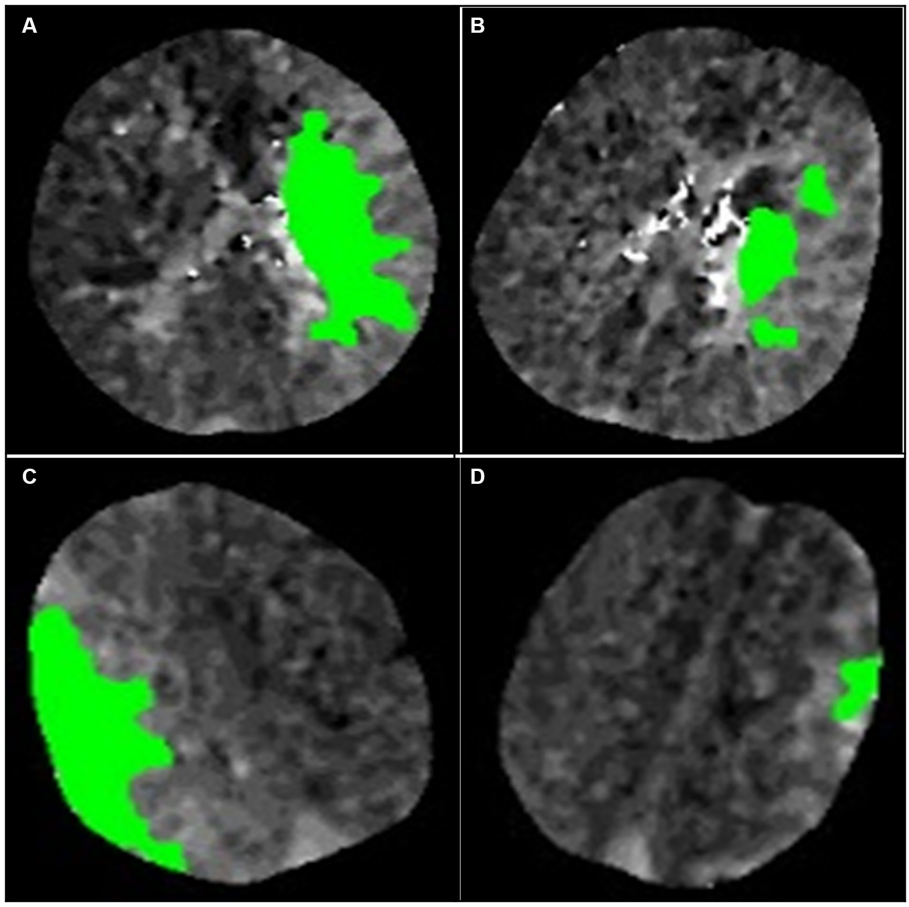


FIGURE 3
Illustration of the definition of the patchy profile and its mimics. **(A,B)** are considered images with patchy profiles. **(C,D)** are considered patchy profile mimics. **(A)** The image is a complete composition, but the edges of the image are extremely irregular and do not satisfy any definition of geometric morphology. **(B)** The image consists of three patches of similar size. **(C)** The image consists of a complete figure but with a smooth curve on one edge. **(D)** The image consists of a patch of minor area.

3.4 Diagnostic performance of the PPS

The PPS was positive in 10 of the 19 patients in the ICAS group but in only 2 of the 32 patients in the intracranial embolism group (52.6% vs. 6.3%, respectively, $p < 0.001$). The sensitivity, specificity, PPV, NPV, and accuracy of PPS for detecting ICAS were 52.6, 93.8, 83.3, 76.9, and 78.4%. The PPS showed a better performance in predicting ICAS (AUC, 0.73; 95% CI: 0.58, 0.89; $p = 0.003$) over HIR ≤ 0.22 (AUC, 0.65; 95% CI: 0.48, 0.81; $p = 0.08$) and CBV ≥ 0.90 (AUC, 0.55; 95% CI: 0.39, 0.72; $p = 0.53$). The ROC curves are shown in Figure 6.

4 Discussion

In the study, an effortless and time-saving imaging technique was discovered, which facilitates the differentiation between intracranial atherosclerotic stenosis and intracranial embolism in acute anterior circulation stroke patients with MCA M1. The study found that the presence of PPS is more likely in patients with ICAS-related LVO. The existing literature on the prognostic capacity of AI software images for stroke etiology is limited. This study could contribute to facilitating further research on AI software imaging across a diverse range of clinical practitioners.

4.1 Potential mechanisms underlying PPS

The PPS may arise because the collateral circulation of patients with atherosclerotic stenosis is better than that of patients with embolism due to the chronic and usually slow progression, allowing sufficient time for opening and formation of collateral circulation (24–26). While the CTA collateral score did not reach statistical significance in this study, there was an observable proportional disparity between the positive and negative groups (22). The proportion of patients with good collateral circulation was significantly higher in the PPS-positive group compared with the PPS-negative group, whereas the proportion of patients with poor collateral circulation was significantly lower. Such findings potentially stem from the limited sample size inherent to our study and disparities in the methodologies employed for collateral circulation assessment. Consequently, a comprehensive large sample study is warranted to validate and substantiate our hypothesis. Chronic cerebral hypoperfusion may promote the formation of intracranial collateral

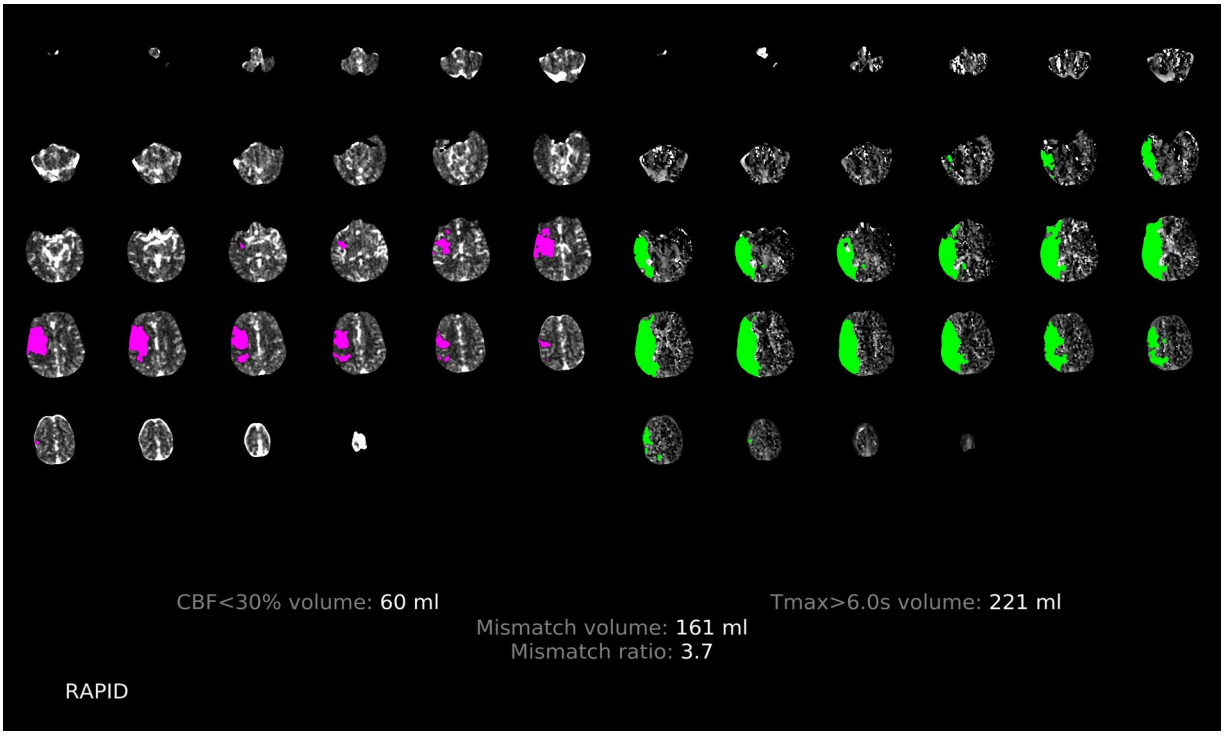


FIGURE 4
Illustration of the patchy profile sign negative. A male patient aged 68 years with occlusion of the right middle cerebral artery M1, where images with the patchy profile sign accounted for less than half of all images with hypoperfusion.

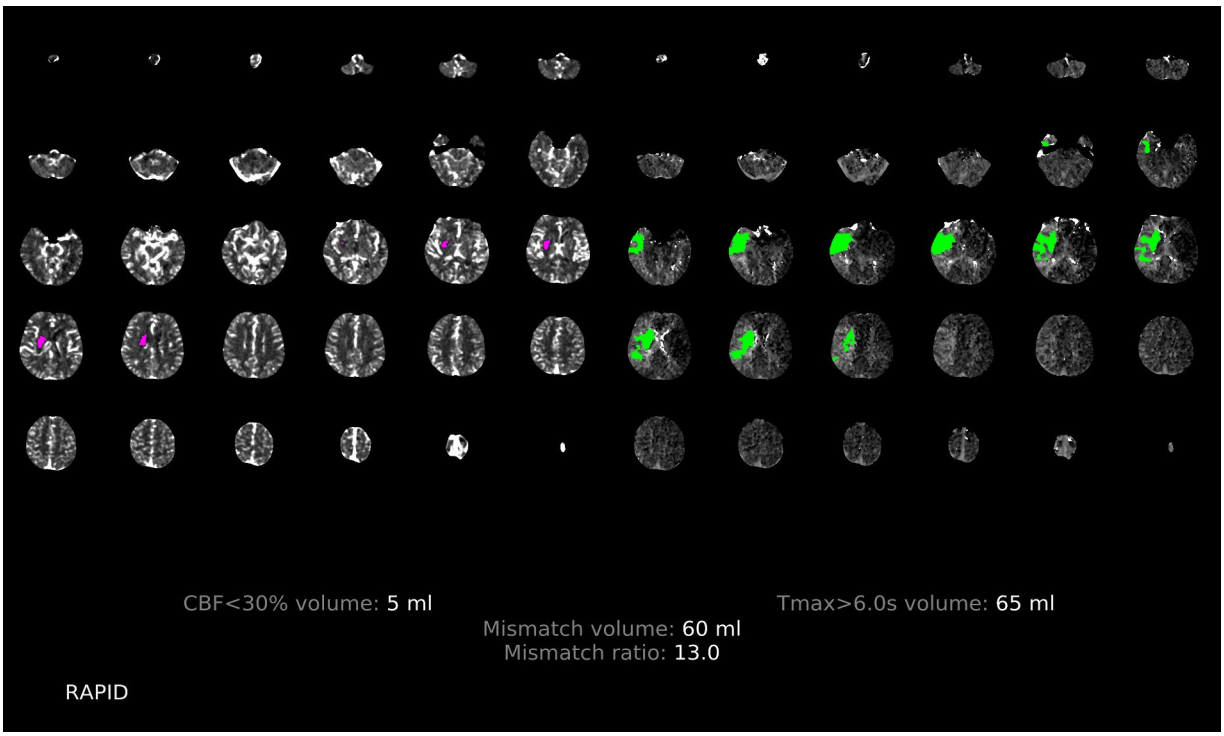


FIGURE 5
Illustration of the patchy profile sign positive. A male patient aged 66 years with occlusion of the right middle cerebral artery M1, where images with the patchy profile sign accounted for more than half of all images with hypoperfusion.

TABLE 1 Clinical, Demographic, and Radiological Characteristics of Patients with and without Patchy Profile Sign.

Variables	Patchy profile sign positive ($n = 12$)	Patchy profile sign negative ($n = 39$)	P value
Male, n (%)	8 (66.7)	20 (51.3)	0.35
Age, years	69.1 ± 12.7	68.9 ± 11.5	0.97
Smoking, n (%)	5 (41.7)	11 (28.2)	0.48
Drinking, n (%)	4 (33.3)	11 (28.2)	0.73
Hypertension, n (%)	7 (58.3)	16 (41.0)	0.29
Diabetes, n (%)	2 (16.7)	11 (28.2)	0.71
Hyperlipidemia, n (%)	0 (0.0)	14 (35.9)	0.02
Onset to imaging time (min)	210.0 (158.0, 480.0)	240.0 (150.0, 420.0)	0.70
Onset to puncture time (min)	300.0 (240.0, 713.0)	390.0 (270.0, 720.0)	0.52
Atrial fibrillation, n (%)	4 (33.3)	18 (46.2)	0.43
Intravenous thrombolysis, n (%)	4 (33.3)	9 (23.1)	0.47
Systolic blood pressure, (mmHg)	147.0 ± 15.4	145.0 ± 28.4	0.82
Diastolic blood pressure (mmHg)	84.2 ± 10.5	79.9 ± 16.2	0.39
Random intravenous blood glucose (mmol/L)	5.5 (5.3, 6.9)	6.9 (5.7, 10.8)	**0.046**
Admission NIHSS scores	12.0 (8.0, 19.0)	16.0 (12.0, 21.0)	0.051
24 h NIHSS	8.4 ± 7.4	11.2 ± 7.1	0.25
GCS	15.0 (12.0, 15.0)	13.0 (10.0, 15.0)	0.06
CBF < 30%, (mL)	0.0 (0.0, 0.0)	15.00 (0.0, 48.0)	**0.001**
Tmax>6 s, (mL)	61.5 ± 38.2	158.8 ± 50.2	**<0.001**
Mismatch ratio[#]	13.9(13.0, 14.7)	5.5 (2.8, 12.0)	0.13
HIR	0.1 (0, 0.2)	0.5 (0.4, 0.6)	**<0.001**
CBV	0.8 (0.7, 1.0)	0.7 (0.6, 0.8)	**0.01**
CTA collateral score			0.07
Good (3)	6 (50. 0)	7 (17. 9)	
Intermediate (2)	5 (41. 7)	23 (60. 0)	
Poor (0–1)	1 (8.3)	9 (23. 1)	
ASPECTS	8.5 (7.0,9.8)	7.0 (6.0, 9.0)	0.10
Mismatch volume, (mL)*	60.6 ± 37.8	128.4 ± 44.3	**<0.001**

[#]Mismatch ratio is defined as Tmax > 6 s divided by CBF < 30%. *Mismatch volume refers to the regions where Tmax > 6 s does not match CBF < 30%. GCS, Glasgow Coma Scale; CBV, cerebral blood volume; HIR, hypoperfusion intensity ratio.

arteries, as demonstrated in an animal model of bilateral carotid artery occlusion (27). Furthermore, the establishment of intracranial collateral circulation is related to the severity and velocity of cerebral vascular stenosis (28). Additionally, a previous study demonstrated that chronic cerebrovascular stenosis induces prolonged cerebral hypoperfusion, inducing a hypoxic-tolerant state in brain tissue. This results in an increased concentration of vascular growth factors, promoting the establishment of collateral circulation, and providing protection against cerebral ischemia (29). Furthermore, it was found that some patients who underwent endovascular treatment and achieved complete recanalization still experienced considerable hypoperfusion on rapid perfusion analysis. The mechanism behind this phenomenon remains unclear and may be related to inadequate microcirculatory reperfusion (30, 31). ICAS-related LVO with a more robust collateral circulation may encounter improved microcirculatory reperfusion after occlusion, leading to patchy areas of hypoperfusion. In contrast, embolization-related LVO may experience inadequate microcirculatory reperfusion due to the rapid occlusion, frequently exhibiting areas of hypoperfusion nearly identical to the region supplied by the responsible vessel. Our hypothesis posits that the presence of PPS in the hypoperfusion region may be attributed to the superior microcirculatory reperfusion in ICAS-related LVO compared to embolization-related LVO. Naturally, the confirmation of specific pathophysiological mechanisms requires further studies.

4.2 Imaging characteristics of ICAS-related LVO vs. embolism-related LVO

The study found that ICAS-related LVO had smaller areas of hypoperfusion compared to embolization-related LVO, which is consistent with prior investigations (22, 32). Contrary to embolization-related LVO, EVT for ICAS-related LVO is associated with a higher incidence of intraoperative reocclusion, an extended

TABLE 2 Clinical, demographic, and radiological characteristics of patients with ICAS and intracranial embolism.

Variables	The ICAS (n = 19)	The IE (n = 32)	P value
Male, n (%)	15 (78.9)	13 (40.6)	**0.008**
Age, years	65.5 ± 11.6	71.0 ± 11.4	0.10
Smoking, n (%)	10 (52.6)	6 (18.8)	**0.01**
Drinking, n (%)	8 (42.1)	7 (21.9)	0.13
Hypertension, n (%)	11 (57.9)	12 (37.5)	0.16
Diabetes, n (%)	5 (26.3)	8 (25.0)	1.00
Hyperlipidemia, n (%)	3 (15.8)	11 (34.4)	0.15
Onset to imaging time (min)	240.0 (180.0, 600.0)	240.0 (127.5, 382.5)	0.12
Onset to puncture time (min)	390.0 (280.0, 990.0)	360.0 (232.5, 585.0)	0.14
Atrial fibrillation, n (%)	2 (10.5)	20 (62.5)	**<0.001**
Intravenous thrombolysis, n (%)	5 (26.3)	8 (25.0)	1.00
Systolic blood pressure, (mmHg)	154.2 ± 26.8	140.3 ± 24.2	0.06
Diastolic blood pressure, (mmHg)	85.0 ± 18.6	78.5 ± 12.3	0.14
Random intravenous blood glucose (mmol/L)	5.9 (5.3, 9.2)	6.7 (5.6, 10.7)	0.28
Admission NIHSS scores	12.0 (8.0, 19.0)	17.0 (13.0, 21.8)	**0.003**
24 h NIHSS	9.2 ± 6.5	11.3 ± 7.6	0.30
GCS	15.0 (11.0, 15.0)	12.5 (10.0, 15.0)	0.12
CBF < 30%, (mL)	0.0 (0.0, 5.0)	15.0 (0.0, 6.0)	**0.002**
Tmax>6 s, (mL)	105.3 ± 67.2	154.1 ± 53.6	**0.006**
Mismatch ratio[#]	12.7 (6.7, 18.2)	5.2 (2.6, 11.7)	0.05
HIR	0.3 (0.0, 0.4)	0.5 (0.4, 0.6)	**0.001**
CBV	0.8 (0.8, 0.9)	0.7 (0.6, 0.8)	**0.02**
CTA collateral score			0.71
Good (3)	6 (31.6)	7 (21.9)	
Intermediate (2)	10 (52.6)	18 (56.3)	
Poor (0–1)	3 (15.8)	7 (21.9)	
ASPECTS	9.0 (7.0, 10.0)	7.0 (6.0, 8.8)	**0.005**
Mismatch volume, (mL)*	100.7 ± 64.6	119.4 ± 41.5	0.27
Patchy Profile Sign	10 (52.6)	2 (6.3)	**<0.001**

[#]Mismatch ratio is defined as Tmax > 6 s divided by CBF < 30%. *Mismatch volume refers to the regions where Tmax > 6 s does not match CBF < 30%. ICAS = intracranial atherosclerotic stenosis; CBV = cerebral blood volume; GCS = Glasgow Coma Scale; HIR = hypoperfusion intensity ratio.

operating time, elevated mortality rates, and reduced revascularization rates (33). In the management of ICAS-related LVO, interventions such as stenting, angioplasty, or a combination of these techniques are typically employed to achieve successful revascularization. Accurate identification of potential ICAS and the development of an optimal strategy for ICAS-related LVO ensure more effective recanalization of occluded vessels and a favorable prognosis for patients with this condition (34, 35). In this study, patients with ICAS were more likely to be smokers but exhibited lower rates of atrial fibrillation, aligning with the findings of a separate study (24). The results of this study indicate that patients with ICAS, in comparison to those with intracranial embolism, manifested a lower NIHSS score, potentially attributed to superior collateral circulation in ICAS patients (26). In addition, the hypoperfusion volume, ischemic core volume, mismatch volume and HIR index in patients with intracranial embolism surpassed those in ICAS patients, while the CBV index in the embolism group was lower than that in the ICAS group. This finding aligns with previous studies (22, 24), and may be associated with the sudden onset, insufficient collateral circulation, and the presence of a large thrombus in patients with intracranial embolism (36).

4.3 Challenges in identifying etiology of AIS in emergency settings

It can be challenging to differentiate accurately between these two causes in clinical settings, particularly during emergency scenarios (37). The following imaging techniques are considered beneficial in assessing the nature of the lesions. High-resolution vessel wall MRI could help to identify ICAS (36, 38), while MRI examinations require significant time, patient cooperation, and are impractical for preoperatively assessing emergency surgery

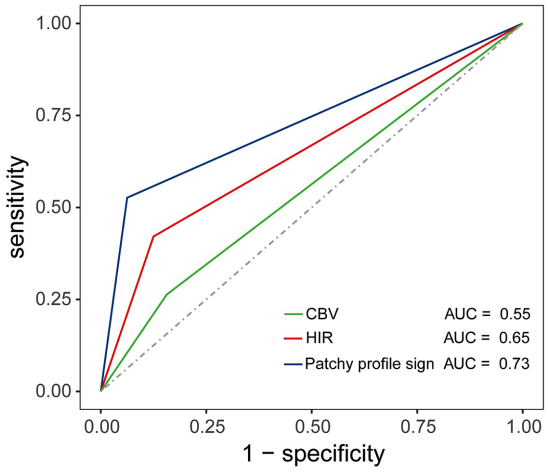


FIGURE 6
Illustration of the receiver operating characteristic (ROC) curve depicting the diagnostic value of HIR, CBV, and the patchy profile sign in predicting ICAS. The area under the curve (AUC) confirmed the superior value of the patchy profile sign in predicting ICAS (AUC, 0.73; $p = 0.003$) compared to HIR (AUC, 0.65; $p = 0.08$) and CBV (AUC, 0.55; $p = 0.53$).

candidates. Kim et al. reported that patients positive for hyperdense middle cerebral artery sign (HMCAS) had a higher incidence of intracranial embolism (67.8% vs. 48.9%, $p = 0.005$), whereas HMCAS-negative patients exhibited a higher incidence of ICAS (31.9% vs. 12.7%, $p = 0.001$) (39). However, several studies have demonstrated that HMCAS does not exhibit a significant correlation with the etiology of AIS (40, 41). Currently, identifying the etiology of AIS using HMCAS still requires extensive observational and demonstrative research. Furthermore, ICAS can be identified by the microcatheter first-pass effect observed during digital subtraction angiography (DSA) owing to the low burden of fresh thrombosis (14). However, the inability to clarify the etiology of the disease before surgery limits the promotion of clinical applications. A recent study found that a Tmax ratio of >10 s/>6 s could predict ICAS-related LVO with or without embolic sources before EVT (42). However, the study did not provide a specific cut-off value, and some patients could not calculate the ratio due to a Tmax >6 s value of zero. Therefore, it is necessary to identify more effective and simpler imaging signs to differentiate ICAS-related from embolization-related LVO. A recent study suggested that HIR ≤ 0.22 (AUC, 0.85; 95% CI: 0.75, 0.96) and CBV ≥ 0.90 (AUC = 0.92, 95% CI: 0.81, 0.98) could serve as valid predictive biomarkers for ICAS (22). In this study, we dichotomized the two indices based on the cut-off values and compared them with the PPS. The analysis revealed that PPS had a higher AUC value for predicting atherosclerosis compared with HIR and CBV. Both HIR and CBV parameters are automatically calculated by software based on perfusion images and are currently used to reflect collateral circulation. A recent investigation has affirmed a significant association between favorable HIR and atherosclerosis (43). Importantly, in our current study, we observed the presence of the PPS in the perfusion images, suggesting a potential shared etiological mechanism with the two indicators.

4.4 Advantages and limitations of the study

Imaging signs such as thrombus imaging, diffusion-weighted imaging, vascular calcification, and collateral circulation are primarily obtained through catheter angiography or thrombectomy, rather than during routine CT scans (44, 45). MRI scanning can also yield some signs, but it necessitates that the patient keeps their heads motionless throughout the examination period. Recognizing these signs requires the expertise of specialized neurologists and imaging physicians, thus limiting their clinical application. A significant advantage of this research is the introduction of a novel imaging approach that can be easily detected in clinical environments. However, our study has some limitations. First, the study was retrospective with a small sample size, highlighting the need for additional prospective studies with larger sample sizes. Second, this study exclusively included patients with occlusion of the MCA M1. Subgroups were identified through imaging, and patients with persistent vascular occlusion were excluded, leaving uncertainty about whether similar conditions exist in patients with occlusion of other vessels. Third, while congestive heart failure in some patients may lead to decreased cerebral perfusion, it remains unclear whether these patients exhibit differences in AI software perfusion images. Consequently, such patients were not excluded from the present study, potentially introducing selection bias into the results. Furthermore, the imaging analysis in this study was conducted using the RAPID software, which might potentially limit its widespread applicability. Several software options are available for perfusion calculations, and further investigation is necessary to determine whether alternative software can produce comparable outcomes. Further research is required for the differential analysis of imaging features in patients with AIS of various etiologies. This will contribute to a more comprehensive assessment of AIS etiology, guiding the selection of clinical treatment options.

5 Conclusion

In conclusion, this study introduces a novel perfusion image sign associated with ICAS. The PPS may function as a specific imaging marker for the identification of ICAS and could potentially guide subsequent endovascular revascularization therapy. Further confirmation through prospective studies with larger sample sizes is necessary to validate the findings of this study.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Yongchuan Hospital of Chongqing Medical University Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for

participation from the participants or the participants' legal guardians/next of kin because of the retrospective nature of the study.

Author contributions

LinZ: Conceptualization, Data curation, Formal analysis, Writing – original draft. HX: Data curation, Writing – original draft. XB: Data curation, Writing – review & editing. JL: Methodology, Writing – review & editing. GT: Data curation, Writing – original draft. YC: Project administration, Writing – original draft. LibZ: Project administration, Writing – review & editing. DY: Conceptualization, Writing – review & editing. LL: Project administration, Supervision, Writing – review & editing. SL: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The study was supported by the National Natural Science Foundation of China [grant no. 81803309]; Chongqing Science and Health Joint Medical Research Project [grant no. 2019QNXM014]; and the Public Welfare Science and Technology Projects of Yongchuan District in 2022 [grant no. 2022yc-jckx20053].

Acknowledgments

We thank Professor Qi Li for his contributions to the creation of the paper.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Berkhemer OA, Fransen PS, Beumer D, van den Berg LA, Lingsma HF, Yoo AJ, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *N Engl J Med*. (2015) 372:11–20. doi: 10.1056/NEJMoa1411587

2. Goyal M, Demchuk AM, Menon BK, Eesa M, Rempel JL, Thornton J, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med*. (2015) 372:1019–30. doi: 10.1056/NEJMoa1414905

3. Saver JL, Goyal M, Bonafe A, Diener HC, Levy EI, Pereira VM, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *N Engl J Med*. (2015) 372:2285–95. doi: 10.1056/NEJMoa1415061

4. Campbell BC, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med*. (2015) 372:1009–18. doi: 10.1056/NEJMoa1414792

5. Jovin TG, Chamorro A, Cobo E, de Miquel MA, Molina CA, Rovira A, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N Engl J Med*. (2015) 372:2296–306. doi: 10.1056/NEJMoa1503780

6. Zi W, Qiu Z, Wu D, Li F, Liu H, Liu W, et al. Assessment of endovascular treatment for acute basilar artery occlusion via a Nationwide prospective registry. *JAMA Neurol*. (2020) 77:561–73. doi: 10.1001/jamaneurol.2020.0156

7. Lapergue B, Blanc R, Gory B, Labreuche J, Duhamel A, Marnat G, et al. Effect of endovascular contact aspiration vs stent retriever on revascularization in patients with acute ischemic stroke and large vessel occlusion: the ASTER randomized clinical trial. *JAMA*. (2017) 318:443–52. doi: 10.1001/jama.2017.9644

8. Wang Y, Zhao X, Liu L, Soo YO, Pu Y, Pan Y, et al. Prevalence and outcomes of symptomatic intracranial large artery stenoses and occlusions in China: the Chinese intracranial atherosclerosis (CICAS) study. *Stroke*. (2014) 45:663–9. doi: 10.1161/strokeaha.113.003508

9. Topcuoglu MA, Liu L, Kim DE, Gurol ME. Updates on prevention of Cardioembolic strokes. *J Stroke*. (2018) 20:180–96. doi: 10.5853/jos.2018.00780

10. Khatri P, Yeatts SD, Mazighi M, Broderick JP, Liebeskind DS, Demchuk AM, et al. Time to angiographic reperfusion and clinical outcome after acute ischaemic stroke: an analysis of data from the interventional Management of Stroke (IMS III) phase 3 trial. *Lancet Neurol*. (2014) 13:567–74. doi: 10.1016/s1474-4422(14)70066-3

11. Bang OY. Intracranial atherosclerosis: current understanding and perspectives. *J Stroke*. (2014) 16:27–35. doi: 10.5853/jos.2014.16.1.27

12. Baek JH, Kim BM, Heo JH, Kim DJ, Nam HS, Kim YD. Outcomes of endovascular treatment for acute intracranial atherosclerosis-related large vessel occlusion. *Stroke*. (2018) 49:2699–705. doi: 10.1161/strokeaha.118.022327

13. Yi TY, Chen WH, Wu YM, Zhang MF, Zhan AL, Chen YH, et al. Microcatheter "first-pass effect" predicts acute intracranial artery atherosclerotic disease-related occlusion. *Neurosurgery*. (2019) 84:1296–305. doi: 10.1093/neuros/nyy183

14. Al Kasab S, Almadidy Z, Spiotta AM, Turk AS, Chaudry MI, Hungerford JP, et al. Endovascular treatment for AIS with underlying ICAD. *J Neurointerv Surg*. (2017) 9:948–51. doi: 10.1136/neurintsurg-2016-012529

15. Jia B, Feng L, Liebeskind DS, Huo X, Gao F, Ma N, et al. Mechanical thrombectomy and rescue therapy for intracranial large artery occlusion with underlying atherosclerosis. *J Neurointerv Surg*. (2018) 10:746–50. doi: 10.1136/neurintsurg-2017-013489

16. Peng G, Zhang Y, Miao Z. Incidence and risk factors of in-stent restenosis for symptomatic intracranial atherosclerotic stenosis: a systematic review and Meta-analysis. *AJNR Am J Neuroradiol*. (2020) 41:1447–52. doi: 10.3174/ajnr.A6689

17. Warach SJ, Luby M, Albers GW, Bammer R, Bivard A, Campbell BC, et al. Acute stroke imaging research roadmap III imaging selection and outcomes in acute stroke reperfusion clinical trials: consensus recommendations and further research priorities. *Stroke*. (2016) 47:1389–98. doi: 10.1161/strokeaha.115.012364

18. Potreck A, Scheidecker E, Weyland CS, Neuberger U, Herweh C, Möhlenbruch MA, et al. RAPID CT perfusion-based relative CBF identifies good collateral status better than Hypoperfusion intensity ratio, CBV-index, and time-to-maximum in anterior circulation stroke. *AJNR Am J Neuroradiol*. (2022) 43:960–5. doi: 10.3174/ajnr.A7542

19. Albers GW, Marks MP, Kemp S, Christensen S, Tsai JP, Ortega-Gutierrez S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med*. (2018) 378:708–18. doi: 10.1056/NEJMoa1713973

20. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med*. (2018) 378:11–21. doi: 10.1056/NEJMoa1706442

21. Olivot JM, Mlynash M, Inoue M, Marks MP, Wheeler HM, Kemp S, et al. Hypoperfusion intensity ratio predicts infarct progression and functional outcome in the DEFUSE 2 cohort. *Stroke*. (2014) 45:1018–23. doi: 10.1161/strokeaha.113.003857

22. Imaoka Y, Shindo S, Miura M, Terasaki T, Mukasa A, Todaka T. Hypoperfusion intensity ratio and CBV index as predictive parameters to identify underlying intracranial atherosclerotic stenosis in endovascular thrombectomy. *J Neuroradiol*. (2023) 50:424–30. doi: 10.1016/j.neurad.2022.10.005

23. Tan JC, Dillon WP, Liu S, Adler F, Smith WS, Wintermark M. Systematic comparison of perfusion-CT and CT-angiography in acute stroke patients. *Ann Neurol*. (2007) 61:533–43. doi: 10.1002/ana.21130

24. Lee JS, Hong JM, Lee KS, Suh HI, Demchuk AM, Hwang YH, et al. Endovascular therapy of cerebral arterial occlusions: intracranial atherosclerosis versus embolism. *J Stroke Cerebrovasc Dis*. (2015) 24:2074–80. doi: 10.1016/j.jstrokecerebrovasdis.2015.05.003

25. Zhang X, Luo G, Jia B, Mo D, Ma N, Gao F, et al. Differences in characteristics and outcomes after endovascular therapy: a single-center analysis of patients with

vertebrobasilar occlusion due to underlying intracranial atherosclerosis disease and embolism. *Interv Neuroradiol*. (2019) 25:254–60. doi: 10.1177/1591019918811800

26. Guglielmi V, LeCouffe NE, Zinkstok SM, Compagne KCJ, Eker R, Treurniet KM, et al. Collateral circulation and outcome in atherosclerotic versus Cardioembolic cerebral large vessel occlusion. *Stroke*. (2019) 50:3360–8. doi: 10.1161/strokeaha.119.026299

27. Jing Z, Shi C, Zhu L, Xiang Y, Chen P, Xiong Z, et al. Chronic cerebral hypoperfusion induces vascular plasticity and hemodynamics but also neuronal degeneration and cognitive impairment. *J Cereb Blood Flow Metab*. (2015) 35:1249–59. doi: 10.1038/jcbfm.2015.55

28. Liebeskind DS, Cotsonis GA, Saver JL, Lynn MJ, Cloft HJ, Chimowitz MI. Collateral circulation in symptomatic intracranial atherosclerosis. *J Cereb Blood Flow Metab*. (2011) 31:1293–301. doi: 10.1038/jcbfm.2010.224

29. Kitagawa K, Yagita Y, Sasaki T, Sugiura S, Omura-Matsuoka E, Mabuchi T, et al. Chronic mild reduction of cerebral perfusion pressure induces ischemic tolerance in focal cerebral ischemia. *Stroke*. (2005) 36:2270–4. doi: 10.1161/01.STR.0000181075.77897.0e

30. Kosior JC, Buck B, Wannamaker R, Kate M, Liapounova NA, Rempel JL, et al. Exploring reperfusion following endovascular Thrombectomy. *Stroke*. (2019) 50:2389–95. doi: 10.1161/strokeaha.119.025537

31. Rubiera M, Garcia-Tornel A, Olivé-Gadea M, Campos D, Requena M, Vert C, et al. Computed tomography perfusion after Thrombectomy: an immediate surrogate marker of outcome after recanalization in acute stroke. *Stroke*. (2020) 51:1736–42. doi: 10.1161/strokeaha.120.029212

32. Yoshimoto T, Inoue M, Tanaka K, Koge J, Shiozawa M, Kamogawa N, et al. Tmax mismatch ratio to identify intracranial atherosclerotic stenosis-related large-vessel occlusion before endovascular therapy. *J Am Heart Assoc*. (2023) 12:e029899. doi: 10.1161/jaha.123.029899

33. Al Kasab S, Almallouhi E, Alawieh A, Wolfe S, Fargen KM, Arthur AS, et al. Outcomes of rescue endovascular treatment of emergent large vessel occlusion in patients with underlying intracranial atherosclerosis: insights from STAR. *J Am Heart Assoc*. (2021) 10:e020195. doi: 10.1161/jaha.120.020195

34. Park H, Baek JH, Kim BM. Endovascular treatment of acute stroke due to intracranial atherosclerotic stenosis-related large vessel occlusion. *Front Neurol*. (2019) 10:308. doi: 10.3389/fneur.2019.00308

35. Kim JH, Jung YJ, Chang CH. Feasibility and safety of the strategy of first stenting without retrieval using solitaire FR as a treatment for emergent large-vessel occlusion due to underlying intracranial atherosclerosis. *J Neurosurg*. (2021) 135:1091–9. doi: 10.3171/2020.8.Jns202504

36. Dieleman N, van der Kolk AG, Zwanenburg JJ, Harteveld AA, Biessels GJ, Luijten PR, et al. Imaging intracranial vessel wall pathology with magnetic resonance imaging: current prospects and future directions. *Circulation*. (2014) 130:192–201. doi: 10.1161/circulationaha.113.006919

37. Anderson CS, Song L, Liu J. Therapeutic strategies for intracranial atherosclerosis. *JAMA*. (2022) 328:529–31. doi: 10.1001/jama.2022.11525

38. Kim SK, Yoon W, Heo TW, Park MS, Kang HK. Negative susceptibility vessel sign and underlying intracranial atherosclerotic stenosis in acute middle cerebral artery occlusion. *AJNR Am J Neuroradiol*. (2015) 36:1266–71. doi: 10.3174/ajnr.A4280

39. Kim SK, Baek BH, Lee YY, Yoon W. Clinical implications of CT hyperdense artery sign in patients with acute middle cerebral artery occlusion in the era of modern mechanical thrombectomy. *J Neurol*. (2017) 264:2450–6. doi: 10.1007/s00415-017-8655-0

40. Topcuoglu MA, Arsava EM, Kursun O, Akpinar E, Erbil B. The utility of middle cerebral artery clot density and burden assessment by noncontrast computed tomography in acute ischemic stroke patients treated with thrombolysis. *J Stroke Cerebrovasc Dis*. (2014) 23:e85–91. doi: 10.1016/j.jstrokecerebrovasdis.2013.08.026

41. Kuo KH, Chang FC, Lai YJ, Pan YJ. Hyperdense artery sign, clot characteristics, and response to intravenous thrombolysis in Han Chinese people with acute large arterial infarction. *J Stroke Cerebrovasc Dis*. (2016) 25:695–701. doi: 10.1016/j.jstrokecerebrovasdis.2015.11.031

42. Rodrigo-Gisbert M, García-Tornel A, Requena M, Vielba-Gómez I, Bashir S, Rubiera M, et al. Clinico-radiological features of intracranial atherosclerosis-related large vessel occlusion prior to endovascular treatment. *Sci Rep*. (2024) 14:2945. doi: 10.1038/s41598-024-53354-z

43. Ballout AA, Libman RB, Schneider JR, Ayoub MS, Wang JJ, Patsalides A, et al. Hypoperfusion intensity ratio is associated with stroke mechanism in patients undergoing mechanical Thrombectomy. *J Stroke Cerebrovasc Dis*. (2022) 31:106539. doi: 10.1016/j.jstrokecerebrovasdis.2022.106539

44. Suh HI, Hong JM, Lee KS, Han M, Choi JW, Kim JS, et al. Imaging predictors for atherosclerosis-related intracranial large artery occlusions in acute anterior circulation stroke. *J Stroke*. (2016) 18:352–4. doi: 10.5853/jos.2016.00283

45. Kim YW, Hong JM, Park DG, Choi JW, Kang DH, Kim YS, et al. Effect of intracranial atherosclerotic disease on endovascular treatment for patients with acute Vertebrobasilar occlusion. *AJNR Am J Neuroradiol*. (2016) 37:2072–8. doi: 10.3174/ajnr.A4844

frontiers | Frontiers in Neurology

TYPE Study Protocol
PUBLISHED 04 June 2024
DOI 10.3389/fneur.2024.1365986

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Leon Edwards,
Gosford Hospital, Australia
Gergely Feher,
University of Pécs, Hungary

*CORRESPONDENCE
Michal Bar
✉ michal.bar@fno.cz

RECEIVED 05 January 2024
ACCEPTED 25 March 2024
PUBLISHED 04 June 2024

CITATION
Dvorníková K, Kunešová V, Ely M, Ostrý S, Čábal M, Reiser M, Machová L, Pavlínová M, Kondé A, Eliáš P, Jonszta T, Havelka J, Volný O and Bar M (2024) The importance of multimodal CT examination in stroke mimics diagnosis: design of prospective observational multicentre study.
Front. Neurol. 15:1365986.
doi: 10.3389/fneur.2024.1365986

COPYRIGHT
© 2024 Dvorníková, Kunešová, Ely, Ostrý, Čábal, Reiser, Machová, Pavlínová, Kondé, Eliáš, Jonszta, Havelka, Volný and Bar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The importance of multimodal CT examination in stroke mimics diagnosis: design of prospective observational multicentre study

Kateřina Dvorníková[1], Veronika Kunešová[2,3], Marcela Ely[1,3], Svatopluk Ostrý[4], Martin Čábal[1], Martin Reiser[4], Linda Machová[1], Markéta Pavlínová[4], Adéla Kondé[5,6], Pavel Eliáš[2], Tomáš Jonszta[2], Jaroslav Havelka[2], Ondrej Volný[1] and Michal Bar[1]*

[1]Department of Neurology, University Hospital Ostrava, Ostrava, Czechia, [2]Department of Imaging Methods, Faculty of Medicine, Ostrava of University, Ostrava, Czechia, [3]Cerebrovascular research program, International Clinical Research Center, Brno, Czechia, [4]Hospital in České Budejovice A.S., České Budějovice, Czechia, [5]Department of the Deputy Director for Science, Research and Education, University Hospital in Ostrava, Ostrava, Czechia, [6]Department of Applied Mathematics, Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, Ostrava, Czechia

Introduction: Non-contrast computed tomography (CT) and CT angiography are the gold standard in neuroimaging diagnostics in the case of suspected stroke. CT perfusion (CTP) may play an important role in the diagnosis of stroke mimics (SM), but currently, it is not a standard part of the stroke diagnostic procedure. The project is a multicentre prospective observational clinical research focused on refining the diagnostics of stroke and stroke mimics (SM) in hospital care.

Aim: This study aimed to evaluate the degree of specificity and sensitivity of multimodal CT (NCCT, CTA, and CTP) in the diagnosis of SM versus stroke.

Methodology: In this study, we will include 3,000 patients consecutively admitted to the comprehensive stroke centres with a diagnosis of suspected stroke. On the basis of clinical parameters and the results of multimodal CT and magnetic resonance imaging (MRI), the diagnosis of stroke and SM will be established. To clarify the significance of the use of the multimodal CT scan, the analysis will include a comparison of the blinded results for each imaging scan performed by radiologists and AI technology and a comparison of the initial and final diagnosis of the enrolled patients. Based on our results, we will compare the economic indicators and costs that would be saved by not providing inadequate treatment to patients with SM.

Conclusion: The expected outcome is to present an optimised diagnostic procedure that results in a faster and more accurate diagnosis, thereby eliminating the risk of inadequate treatment in patients with SM.

Clinical trial registration: clinicaltrials.gov, NCT06045455.

KEYWORDS

CT perfusion, multimodal CT, stroke, stroke mimics, MRI

Introduction

Stroke mimics (SM) present a wide variety of neurological disorders that exhibit symptoms similar to those of a proven cerebrovascular accident (CVA—stroke), despite not being a true stroke. Every year, stroke affects approximately 28,000 people across the entire age spectrum in the Czech Republic, out of which one-third of patients succumb to the damage, some patients remain with a permanent deficit, and only a proportion of these patients recover fully from this disease. To achieve a positive outcome for patients with stroke, it is essential to initiate treatment early, which requires the early recognition of symptoms and the diagnosis of stroke. However, a rapid assessment can lead to the unintentional treatment of patients with SM, which is not indicated, is costly, and can lead to complications. A meta-analysis of data from 62,664 patients indicates that the average prevalence of SM is 24.8% (with an interval of 20–50%). The frequency of SM varies and depends on where the diagnosis is conducted and whether the patients are assessed by a physician at the emergency department or by a neurologist (1, 2).

SM can be divided into two main categories: medical SM and functional SM. Medical SM consist of brain disorders (migraine, seizures, posterior reversible encephalopathy syndrome/PRES, epidural/subdural haematomas, hypertensive encephalopathy, or brain tumours) and systemic diseases (electrolyte dysfunction and metabolic/toxic disorders). The frequency of the individual types of SM varies across different studies. The most recent comprehensive review showed the following frequencies of the different types of SM: peripheral vestibular syndrome 23.2%, toxic/metabolic disorders 13.2%, seizures 13%, migraine 7.8%, syncope 6%, mononeuropathy 5%, space-occupying brain lesions 4%, acute confusion 1.9%, dementia 1.2%, and spinal cord lesions 0.7% (3).

CT perfusion (CTP) identifies the ischaemic brain tissue in patients with suspected stroke even when non-contrast head computed tomography (NCCT) is normal (4, 5). CTP can estimate brain regions with a high probability of irreversible infarction (ischaemic core) and areas at risk of infarction (penumbra). CTP is a functional examination of the brain tissue that characterises the state of brain perfusion and provides information on the current functional state of the microcirculation (6). These facts provide very important prognostic information in the cases of successful and unsuccessful acute recanalisation treatment. CTP allows the selection of suitable candidates for recanalisation treatment after 6 h in the case of mechanical thrombectomy (MT) and after 4.5 h in the case of intravenous thrombolysis (IVT), or if the time of stroke onset is unknown. Recent studies have demonstrated the benefit of MT within 6 h from the onset of stroke in patients with a large ischaemic core. The importance of CTP as an indication for acute recanalisation treatment has decreased, but the prognostic and diagnostic roles of CTP is still essential (7). Perfusion imaging abnormalities have also been reported to be useful in differentiating SM from acute ischaemic stroke and may inform treatment decision-making (8–11).

Exclusion of ischaemic stroke and early detection of SM are crucial in reducing the number of unindicated intravenous thrombolysis (IVT) treatments, which is the first choice of treatment for patients with stroke. In the randomised Norwegian Tenecteplase Stroke Trial (NOR-TEST) study that assessed the effect of tenecteplase in ischaemic strokes, IVT was administered to 18% of patients (tenecteplase vs. alteplase) with a diagnosis other than ischaemic stroke (12). Unindicated administration of IVT can cause possible negative implications for patients (extra or intracerebral haemorrhage, orolingual angioedema, etc.), thereby leading to permanent health and social stigma.

The existing literature predominantly consists of a small quantity of case series concerning SM and have often included only patients with seizures or migraines. In reality, the spectrum of SM is much wider. In this respect, the present project is unique when compared with the current published studies.

Overall goals of the project

This project aims to optimise hospital diagnostics for stroke and SM symptoms, focusing on the specificity and sensitivity of multimodal brain imaging compared to conventional imaging. The specific goals of the study include evaluating diagnostic outcomes between blinded radiologists/AI technology and clinical diagnosis as well as comparing economic indicators to identify potential savings from avoiding unnecessary treatments.

Methods

The study employs a prospective multicentre observational design in two Czech comprehensive stroke centres, enrolling an estimated 3,000 patients over 3 years.

The inclusion criteria include acute neurological deficits with symptoms lasting up to 24 h and NIHSS $\geq$0 points (National Institutes of Health Stroke Scale) at enrolment.

The exclusion criteria include age < 18 years, previous severe neurological handicap (mRS = 5), contraindications for CT contrast scan or MRI (pregnancy, severe renal failure, allergy to contrast material, or claustrophobia), and patient non-consent. For the trial flow, see the STARD diagram (Figure 1).

Monitored parameters

The following data will be collected and evaluated in Table 1.

Stroke protocol

Multimodal CT

All patients with acute neurological deficits (stroke, transient ischaemic stroke (TIA), or SM) will undergo neuroimaging diagnostic procedures as per the stroke protocol, i.e., initial multimodal brain CT (NCCT, CTP, and computed tomography angiography (CTA)), followed by follow-up NCCT within 24–36 h.

The basic CTP parameters define perfusion maps, which express the mean transit time, cerebral blood flow, and cerebral blood volume. Currently, the maps defining the so-called maximum time (Tmax) are also essential. The AI software package allows us to determine the extent of irreversible ischaemia (decrease in CBF below 30% of the reference bilateral value) and potentially the tissue that can be rescued, i.e., penumbra (prolongation of Tmax over 6 s).

Type of CT device—CT Force Siemens.

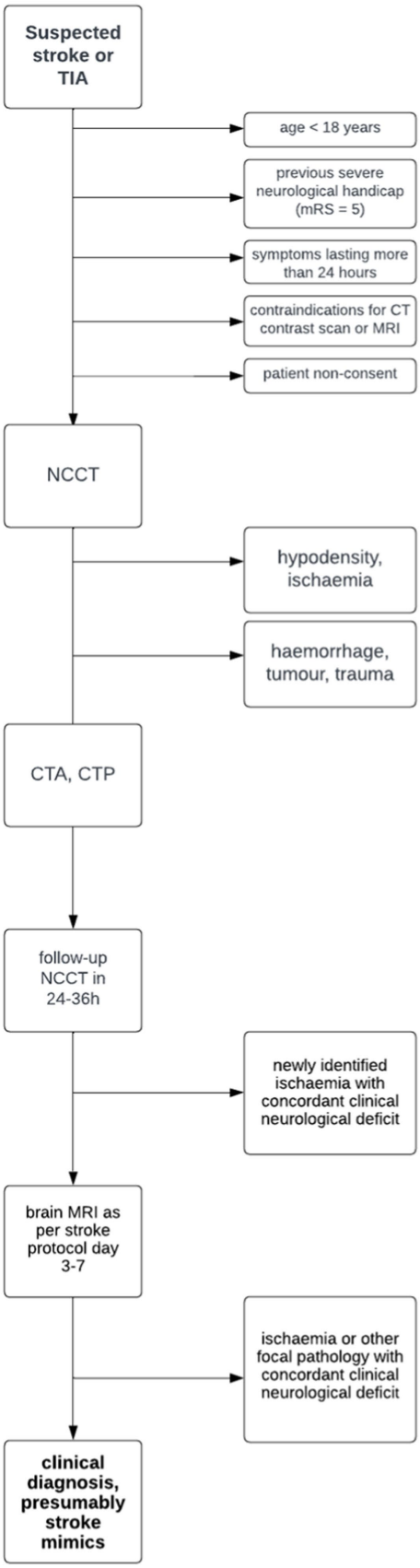


FIGURE 1
STARD diagram.

Stroke protocol: Unenhanced brain: 120 kV (CarekV), 300 mAs (CareDose), collimation 128 × 0.6 mm rotation time 0.5 s Recon: slice 0.75 mm, 5 mm, kernel Hr 40.

CTP: 70 kV, 200 mAs, collimation 48×1.2 mm, brain coverage-4D range: 114 mm.

Cycle time 1.5 s exam time 45 s rotation time 0.25 s contrast medium—40 mL intravenous.

CTA: 120 kV (CarekV), 84 mAs (Care Dose), collimation 128 × 0.6 mm.

Recon: slice 0.5 mm, 3 mm, kernel Hr 36, contrast medium—50 mL intravenous.

MRI examination

All patients without a confirmed concordant hypoperfusion or cerebral ischaemia on their previous CT scans will undergo an MRI examination between the third and seventh day after admission to hospital to confirm the diagnosis of SM.

The MRI examination will be performed using the Siemens Prisma, 3.0 T scanner (Siemens, Erlangen, Germany). The imaging protocol includes:

- localisers;
- diffusion weighted imaging (DWI) with b-factor values 0, 400, and 800 s/mm2, calculated b-value of 1,000 s/mm2. ST 3.0 mm, GAP 0.6 mm, TR 2900 ms, TE 60 ms, FOV 256 mm, BW 868 Hz/Px, number of slices 40, TA 2:09;
- reconstructed maps of an apparent diffusion coefficient (ADC maps);
- 3D FLAIR sequence. The examination will be performed in the sagittal plane with 1 mm MPR axial and coronal reconstructions. TR 5500 ms, TE 383 ms, ST 1 mm, number of slices 160, GAP 0, TI 1800 ms, FA T2 Var. deg., BW 751 Hz/Px, TF 220, TA—acquisition time 3:46. Total examination time 6:12+ shimming time.

Follow-up brain MRI will be performed on the following patients as per the stroke protocol:

1 Patients with a negative initial multimodal CT scan.
2 Patients with discordant CTP results, which means that the location of the hypoperfusion does not correspond with the neurological symptoms, or patients with CTP technical artefacts.
3 Patients with a negative follow-up brain NCCT.
4 Patients with an NIHSS score of zero upon admission to the hospital and 24 h after the onset of TIA.

Analyse of CT and MRI examination

The results of NCCT, CTA, and CTP will be evaluated by a blinded radiologist against the clinical diagnosis and will then be compared with the multimodal CT examination evaluated by AI. The concordance of a positive CTP finding in relation to the clinic will be assessed by a specialist stroke neurologist. The AI software (e-Stroke, Brainomix, Oxford) provides a fully automated scoring tool

TABLE 1 Following data will be collected and evaluated.

Patient history	
Age Sex Comorbidities Long-term medication	Hypertension, diabetes mellitus, heart disease, arrhythmia, hyperlipidemia, mental disorders, cephalgia, migraine, epilepsy, oncological disease, cognitive deficit antithrombotic medication, hormonal anticonception
Hospitalisation data	
Date and time of stroke/stroke mimics Date of hospital admission Hospitalisation, duration	standard, ICU
Systolic/diastolic blood pressure Laboratory parameters-INR, glycemia, Increase in liver tests, kidney failure, drugs in the blood	
Clinical data	
mRS NIHSS Neurological deficit Type of treatment Multimodal brain CT scan on admission Follow-up NCCT Brain MRI results Final diagnosis	Pre-morbid, at discharge on admission, 24 h after the admission to hospital NCCT, CTA, CTP during 24–36 h after the admission in indicated patients stroke, stroke mimics
Follow-up visit	
mRS value Recurrence of symptoms Clinical condition	

for analysing CT and MRI examinations. The software calculates e-ASPECT and the volume of ischaemia; it analyses CTP and the values of collateral circulation. Furthermore, it can determine the area of vascular occlusion and the size of ischaemia on MRI. In this respect, Brainomix e-STROKE software is a unique and irreplaceable tool.

Determination of the definitive diagnosis

Two neurologists specialised in treating and diagnosing stroke will confirm the diagnosis independently.

The following patients will be diagnosed with an ischaemic stroke:

a Patients with a finding of hypodensity corresponding with subacute ischaemia on the initial brain CT examination.
b Patients with a new subacute ischaemic lesion on the follow-up NCCT scan within 36 h from the onset of the neurological deficit.
c Patients with an ischaemic brain lesion on their MRI scan within 3–7 days from the onset of the neurological deficit.

The diagnosis of SM will be determined by the following parameters:

a organic lesion of aetiology other than ischaemic stroke on the NCCT scan,
b negative CTP, discordant CTP, and negative control brain MRI,
c two neurologists and two radiologists specialised in treating and diagnosing stroke will confirm the diagnosis of SM independently, considering the correlation of imaging and clinical data (see clinical data checklist). The radiologists will also assess other thresholds during the CTP examination than the classically established ones (e.g., T max less than 6 s). Both stroke consultants will assess the final diagnosis for the whole group with and without knowledge of the CTP result.

Follow-up visit

Telephone or outpatient appointment within 90 days from the onset of the neurological deficit. During this visit, the patient's mRS value, recurrence of symptoms, and clinical condition should be evaluated.

Quality of life (QoL) evaluation and pharmacoeconomic analysis

We have incorporated a quality of life (QoL) evaluation framework into our study to demonstrate our commitment to a holistic approach to patient care and to enhance the understanding of the benefits of accurate diagnosis beyond the immediate clinical outcome. The QoL framework in our study includes both the

Modified Rankin Scale (mRS) and the quality-adjusted life years (QALY) metrics (13).

The mRS is known for its direct relevance to stroke outcomes and its ability to measure changes in functional status. Furthermore, it is simple and easy-to-use tool that is widely accepted and validated in stroke research. In the context of pharmacoeconomic analysis in our study, the mRS provides a quantifiable outcome measure that can be related to QALYs and other economic evaluations. QALYs are commonly used in health economics. Our study aims to highlight how an accurate diagnosis can lead to a better patient outcome and potentially lower healthcare costs over time (14, 15).

Statistic analysis

A minimum of 718 patients is required to demonstrate significant differences between diagnostic approaches. The analysis involves appropriate statistical tests and logistic regression to assess factors influencing clinical diagnosis. It is expected that the total number of enrolled patients will be 3,000 in order to ensure the availability of all established parameters and, particularly, the relevance of the data obtained. The sample size was determined using the proportional data from the pilot project ($p = 0.34$ for the observation group and a desired effect of at least 10%), 5% significance level, and 80% power.

Numerical variables are presented as medians and ranges or interquartile ranges. Categorical variables are presented as absolute and relative frequencies. The association between demographic, clinical, or radiographic variables and the clinical diagnosis (stroke and stroke mimic) is analysed using appropriate statistical tests—Mann–Whitney test, chi-squared test of independence for contingency tables, or other statistical tests, if necessary. Eventually, logistic regression can be used for analysing the effect of multiple factors on the clinical diagnosis.

The diagnostic accuracy of applied diagnostic methods will be evaluated using common measurements, including accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. The characteristics will be reported with their confidence intervals. The significance level is set to 5%, and statistical analyses will be performed using appropriate statistical software (R, STATA) with maximum available data.

The pharmacoeconomic analysis will include a health technology assessment (HTA) to calculate the potential value of innovation at an early stage, analysing not only the QoL of patients but also the costs associated with the treatment of stroke and SM. The calculation of costs is based on the average cost of thrombolysis and its associated treatment, representing the incremental cost (IC). Economic evaluation is organised as a test of dominance of improved results (meaning fewer complications associated with the administration of thrombolysis) and lower costs.

To calculate the overall benefit, the value of the QoL and the specific parameter of the quality-adjusted life years (QALY) set for the Czech Republic will be used.

The results will be reported as expected ICs and effects (E) expressed in quality-adjusted years of life (QALY).

Pilot project

The design of the proposed research project is based on the pilot analysis that was carried out in the Comprehensive Stroke Centre in the Hospital in Ceske Budejovice, where data from a total of 109 patients with suspected stroke were evaluated. These patients underwent multimodal CT examinations (NCCT, multiphase CTA, and CTP) as per the stroke protocol. MRI examination DWI/FLAIR mismatch was selectively indicated as determined by the clinical context.

Based on the performed examinations, stroke was detected in 72 cases, and SM were detected in 37 cases (34%) (see Tables 2, 3).

Seven patients (6.4%) had symptoms caused by a metabolic disorder and/or a cardiovascular disorder; five patients (4.6%) had symptoms of seizures; and three patients in each group (2.8%) had symptoms of functional disorder, trauma, and neuropsychiatric manifestations (see Table 2). The SM diagnosis was determined in the cases of drug intoxication in two patients (1.8%) and Bell's palsy, vestibular syndrome, migraine, and spondylogenic symptoms in one patient. An inconclusive or unexplained cause of health complications was found in a total of 10 (9.2%) cases.

TABLE 2 Demographic data of patients in a pilot project.

Suspected stroke		109 patients	
Stroke	66%	Stroke mimics	34%
Male	53%	Male	51%
Female	47%	Female	49%
Age (average value)	71.1 ± 15.5	Age (average')	67.9 ± 15.5

Discussion

This project aims to clarify stroke and SM symptoms in hospital care, potentially modifying care provision. It could substantiate international publications and validate the role of multimodal CT in stroke and SM diagnosis.

It is common knowledge that the examination with the highest sensitivity and specificity for the diagnosis of ischaemic stroke is the MR examination of the brain (16). MR examination has numerous contraindications, and as part of an acute diagnosis, stroke is not routinely used. On the other hand, multimodal CT examinations (including non-contrast head CT (NCCT), head and neck CT angiography (CTA), and CTP) may play a significant role in providing accurate information about brain pathology in a short period of time. NCCT (17) and head and neck CTA, which are commonly used for the detection of large vessel occlusion, remain the gold standard in neuroimaging diagnostics in the case of suspected stroke, whilst CTP may play an important role in the diagnosis of SM. However, CTP is currently not a standard part of the diagnostic procedure in most stroke centres. Most patients with SM exhibited discordant clinical findings when compared to CTP imaging. Normal or entirely discordant Tmax of 6 s abnormalities are more common in patients with SM. Therefore, CTP could be crucial when deciding on the use of intravenous thrombolytic therapy (4). The sensitivity of acute NCCT examination for the diagnosis of ischaemic stroke is low (33%) (5).

The outcome of this project is based on the evaluation that the combination of NCCT, CTA, and CTP is superior to NCCT alone or

TABLE 3 Detection of stroke mimics.

Stroke mimics	37	34.0%
Metabolic and/or cardiovascular disorder	7	6.4%
Seizures	5	4.6%
Functional disorder	3	2.8%
Trauma	3	2.8%
Neuropsychiatric manifestations	3	2.8%
Drug intoxication	2	1.8%
Bell's palsy	1	0.9%
Vestibular syndrome	1	0.9%
Migraine	1	0.9%
Spondylogenic symptom	1	0.9%
Inconclusive or unexplained cause	10	9.2%

the combination of NCCT and CTA in the diagnosis of SM in patients with acute neurological deficits. The final diagnosis of SM will be determined based on a unique combination of knowledge of clinical data and multimodal CT and MRI results.

Conclusion

The project's anticipated outcome is an optimised diagnostic procedure, reducing the risk of inadequate treatment for suspected stroke patients. The applied research may yield health and socioeconomic benefits, improving healthcare provision and reducing costs.

Ethics statement

The studies involving humans were approved by Ethics Committee at the University Hospital in Ostrava, No. 503/2022, 23.6.2022; Ethics Committee at the Hospital in České Budějovice, No. 109/22, 24.6.2022. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KD: Investigation, Methodology, Writing – original draft, Writing – review & editing. VK: Conceptualization, Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review & editing. ME: Conceptualization, Data curation, Software, Writing – review & editing. SO: Conceptualization, Investigation, Methodology, Supervision, Visualization, Writing – review & editing. MČ: Data curation, Investigation, Methodology, Writing – review & editing. MR: Investigation, Methodology, Writing – review & editing, Data curation. LM: Data curation, Investigation, Methodology, Writing – review & editing. MP: Data curation, Investigation, Methodology, Writing – review & editing. AK: Formal analysis, Methodology, Project administration, Writing – review & editing. PE: Conceptualization, Formal analysis, Methodology, Software, Supervision, Writing – review & editing. TJ: Data curation, Investigation, Methodology, Software, Writing – review & editing. JH: Data curation, Investigation, Methodology, Software, Writing – review & editing. OV: Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – review & editing. MB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Czech Health Research Council of Ministry of Health (NU23-04-00336); Conceptual development of research organization of Ministry of Health (FNOs/2022); STROCZECH within CZECRIN Large Research Infrastructure (No. LM2023049).

Conflict of interest

MB received a speaker fee from Brainomix ltd in 2021. SO, MR, and MP were employed by Hospital in České Budejovice A.S.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1365986/full#supplementary-material

References

1. Merino JG, Luby M, Benson RT, Davis LA, Hsia AW, Latour LL, et al. Predictors of acute stroke mimics in 8187 patients referred to a stroke service. *J Stroke Cerebrovasc Dis*. (2013) 22:e397–403. doi: 10.1016/j.jstrokecerebrovasdis.2013.04.018

2. Pohl M, Hesszenberger D, Kapus K, Meszaros J, Feher A, Varadi I, et al. Ischemic stroke mimics: a comprehensive review. *J Clin Neurosci*. (2021) 93:174–82. doi: 10.1016/j.jocn.2021.09.025

3. Buck HB, Akhtar N, Alrohimi A, Khan K, Shuaib A. Stroke mimics: incidence, aetiology, clinical features and treatment. *Ann Med*. (2021) 53:420–36. doi: 10.1080/07853890.2021.1890205

4. Siegler JE, Rosenberg J, Cristancho D, Olsen A, Pulst-Korenberg J, Raab L, et al. Computed tomography perfusion in stroke mimics. *Int J Stroke*. (2020) 15:299–307. doi: 10.1177/1747493019869702

5. Smajlović D, Sinanović O. Sensitivity of the neuroimaging techniques in ischemic stroke. *Med Arh*. (2004) 58:282–4.

6. Van Cauwenberge MGA, Dekeyzer S, Nikoubashman O, Dafotakis M, Wiesmann M. Can perfusion CT unmask postictal stroke mimics? A case-control study of 133 patients. *Neurology*. (2018) 91:1918–27. doi: 10.1212/WNL.0000000000006501

7. Doheim MF, Hagrass AI, Elrefaey M, Al-Bayati AR, Bhatt NR. From therapeutic nihilism to armamentarium: A meta-analysis of randomized clinical trials assessing safety and efficacy of endovascular therapy for acute large ischemic strokes. *Interv Neuroradiol*. (2023). doi: 10.1177/15910199231170681

8. Austein F, Huhndorf M, Meyne J, Laufs H, Jansen O, Lindner T. Advanced CT for diagnosis of seizurerelated stroke mimics. *Eur Radiol*. (2017) 28:1791–800. doi: 10.1007/s00330-017-5174-4

9. Shelly S, Maggio N, Boxer M, Blatt I, Tanne D, Orion D. Computed tomography perfusion maps reveal blood flow dynamics in postictal patients: a novel diagnostic tool. *Isr Med Assoc J*. (2017) 19:553–6.

10. Ridolfi M, Granato A, Polverino P, Furlanis G, Ukmar M, Zorzenon I, et al. Migrainous aura as stroke-mimic: the role of perfusion-computed tomography. *Clin Neurol Neurosurg*. (2018) 166:131–5. doi: 10.1016/j.clineuro.2018.01.032

11. Campbell BC, Weir L, Desmond PM, Tu HTH, Hand PJ, Yan B, et al. CT perfusion improves diagnostic accuracy and confidence in acute ischaemic stroke. *J Neurol Neurosurg Psychiatry*. (2013) 84:613–8. doi: 10.1136/jnnp-2012-303752

12. Logallo , Novotny V, Assmus J, Kvistad CE, Alteheld L, Rønning OM, et al. Tenecteplase versus alteplase for management of acute ischaemic stroke (NOR-TEST): a phase 3, randomised, open-label, blinded endpoint trial. *Lancet Neurol*. (2017) 16:781–8. doi: 10.1016/S1474-4422(17)30253-3

13. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull*. (2010) 96:5–21. doi: 10.1093/bmb/ldq033

14. Cheon S, Li CHY, Jeng JS, Wang JD, Ku LJE. The lifetime burden following stroke: long term impact of stroke on survival and quality of life. *Int J Stroke*. (2023) 18:795–803. doi: 10.1177/17474930231165607

15. Zhang T, Liang Z, Lin T, Cohen DJ, Arrieta A, Wang X, et al. Cost-effectiveness of folic acid therapy for primary prevention of stroke in patients with hypertension. *BMC*. (2022) 20:407. doi: 10.1186/s12916-022-02601-z

16. Hand PJ, Wardlaw JM, Rowat AM, Haisma JA, Lindley RI, Dennis MS. Magnetic resonance brain imaging in patients with acute stroke: feasibility and patient related difficulties. *J Neurol Neurosurg Psychiatry*. (2005) 76:1525–7. doi: 10.1136/jnnp.2005.062539

17. Vilela P. Acute stroke differential diagnosis: stroke mimics. *Eur J Radiol*. (2017) 96:133–44. doi: 10.1016/j.ejrad.2017.05.008

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 10 June 2024
DOI 10.3389/fneur.2024.1385013

Check for updates

OPEN ACCESS

EDITED BY
Xiangzhi Bai,
Beihang University, China

REVIEWED BY
Anwar P. P. Abdul Majeed,
Sunway University, Malaysia
Xiaozhu Liu,
Capital Medical University, China
Bin Yi,
Army Medical University, China

*CORRESPONDENCE
Hongbin Yuan
jfjczyy@aliyun.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 11 February 2024
ACCEPTED 28 May 2024
PUBLISHED 10 June 2024

CITATION
Wei H, Huang X, Zhang Y, Jiang G, Ding R, Deng M, Wei L and Yuan H (2024) Explainable machine learning for predicting neurological outcome in hemorrhagic and ischemic stroke patients in critical care.
Front. Neurol. 15:1385013.
doi: 10.3389/fneur.2024.1385013

COPYRIGHT
© 2024 Wei, Huang, Zhang, Jiang, Ding, Deng, Wei and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Explainable machine learning for predicting neurological outcome in hemorrhagic and ischemic stroke patients in critical care

Huawei Wei[1†], Xingshuai Huang[1†], Yixuan Zhang[1†], Guowei Jiang[1†], Ruifeng Ding[1], Mengqiu Deng[1], Liangtian Wei[2] and Hongbin Yuan[1*]

[1]Department of Anesthesiology, Changzheng Hospital, Second Affiliated Hospital of Naval Medical University, Shanghai, China, [2]Jiangsu Province Key Laboratory of Anesthesiology, Xuzhou Medical University, Xuzhou, China

Aim: The objective of this study is to develop accurate machine learning (ML) models for predicting the neurological status at hospital discharge of critically ill patients with hemorrhagic and ischemic stroke and identify the risk factors associated with the neurological outcome of stroke, thereby providing healthcare professionals with enhanced clinical decision-making guidance.

Materials and methods: Data of stroke patients were extracted from the eICU Collaborative Research Database (eICU-CRD) for training and testing sets and the Medical Information Mart for Intensive Care IV (MIMIC IV) database for external validation. Four machine learning models, namely gradient boosting classifier (GBC), logistic regression (LR), multi-layer perceptron (MLP), and random forest (RF), were used for prediction of neurological outcome. Furthermore, shapley additive explanations (SHAP) algorithm was applied to explain models visually.

Results: A total of 1,216 hemorrhagic stroke patients and 954 ischemic stroke patients from eICU-CRD and 921 hemorrhagic stroke patients 902 ischemic stroke patients from MIMIC IV were included in this study. In the hemorrhagic stroke cohort, the LR model achieved the highest area under curve (AUC) of 0.887 in the test cohort, while in the ischemic stroke cohort, the RF model demonstrated the best performance with an AUC of 0.867 in the test cohort. Further analysis of risk factors was conducted using SHAP analysis and the results of this study were converted into an online prediction tool.

Conclusion: ML models are reliable tools for predicting hemorrhagic and ischemic stroke neurological outcome and have the potential to improve critical care of stroke patients. The summarized risk factors obtained from SHAP enable a more nuanced understanding of the reasoning behind prediction outcomes and the optimization of the treatment strategy.

KEYWORDS

critical care, machine learning, model interpretability, prediction model, stroke

Introduction

Stroke encompasses a set of conditions characterized by the sudden rupture or occlusion of cerebral blood vessels, ultimately resulting in insufficient blood flow and subsequent damage to brain tissue. Clinically, stroke is broadly classified into two main types—ischemic and hemorrhagic—with the latter comprising intracerebral and

subarachnoid hemorrhage forms (1). Stroke affects a staggering one in every four individuals over 25 years of age, rendering it the second most common cause of mortality and third leading cause of disability among adult populations worldwide (2). Approximately 16 million people worldwide suffer from various motor and cognitive impairments as a result of stroke, which are often unavoidable sequelae for stroke patients, and severely affects the mobility and quality of life of stroke victims (3).

Acute stroke patients often enter the intensive care unit (ICU) due to consciousness disorders, cardiopulmonary complications, circulatory instability, or acute thrombolytic therapy (4). Compared with patients admitted to a dedicated neurological ward or stroke unit, those with stroke who are admitted to the ICU exhibit heightened neurological severity, notable impairment of consciousness at a moderate to severe level, often necessitating mechanical ventilation, and encounter an elevated risk of hospital mortality (5, 6). ICU provides complex and resource-intensive treatment for hospitalized patients with severe conditions, but current medical resources are often insufficient to meet the needs of ICU patients, and hospitals face pressure to improve critical care efficiency and reduce costs (7). Early prediction of neurological outcome in critically ill stroke patients can provide important references for patients and their families, and can also guide clinicians to give the best intervention measures to patients.

In contrast to conventional predictive models that rely on established variables for computation, machine learning (ML) approaches offer the distinct advantage of incorporating a broader range of variables that more comprehensively capture the intricacies and inherent unpredictability of human physiology (8, 9). Consequently, ML has emerged as a promising tool in the medical field, with its capacity to integrate abundant variables, extract nuanced insights, and generalize acquired knowledge to novel cases with remarkable efficiency and precision (10, 11). Furthermore, interpretable machine learning is increasingly being applied in clinical research, demonstrating robust clinical applicability and guiding capabilities (12, 13).

In this work, we aimed to construct ML models for early and effective prediction of neurological outcome at hospital discharge in critically ill patients with hemorrhagic and ischemic stroke, and employed the shapley additive explanations (SHAP) methods to elucidate the underlying reasons and decision-making processes involved within the optimal algorithm.

Materials and methods

Study design

The implementation of the study design was shown in Figure 1. The present study was a retrospective modeling study utilizing data from two widely used databases—the eICU Collaborative Research Database (eICU-CRD v2.0) spanning 2014–2015 and the Medical Information Mart for Intensive Care IV (MIMIC IV version 2.2) covering 2008–2019. The author of this study underwent rigorous training, culminating in certification (number 49437998), and was tasked with data extraction following secure access to both databases.

Participants

In this study, patients diagnosed with stroke according to the ninth and 10th revisions of the international classification of diseases were included (Table 1). These patients were then stratified into hemorrhagic and ischemic cohorts for comparative analysis. Inclusion criteria included individuals over 18 years of age but under 89 years of age who had been in the ICU for at least more than 24 h, along with a Glasgow Coma Scale (GCS) score within 24 h of admission and a documented motor GCS score within 24 h prior to discharge. It is important to note that in the case of repeat ICU admissions, only data relating to the first ICU admission were retained.

Variables extraction and outcome

In this study, detailed demographic data were collected on age, gender, race, weight, height, and body mass index (BMI). The maximum, minimum and mean values of vital signs during the initial 24 h of ICU admission were extracted, encompassing heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), mean blood pressure (MBP), temperature, respiratory rate (RR), and oxygen saturation (SpO_2). Laboratory parameters obtained within the first 24 h of admission were also extracted. For certain parameters with multiple measurements, both the maximum and minimum values were evaluated. Moreover, the medical interventions employed during the 1st day of admission, such as mechanical ventilation and renal replacement therapy, along with the illness severity scoring systems, namely Charlson comorbidity index, GCS, acute physiology score III (APS III), and sequential organ failure assessment (SOFA), were recorded. Table 2 summarized the variables extracted.

The objective of this study was to investigate neurological status at the time of hospital discharge. In the stroke population, the National Institute of Health stroke scale (NIHSS) is a widely accepted metric to determine neurological outcomes. However, since NIHSS was not recorded in eICU-CRD or MIMIC IV, we adopted a surrogate neurological outcome marker based on the motor subscore of the Glasgow Coma Score (mGCS) at discharge. The mGCS score was stratified into two categories, favorable (mGCS score of 6) and unfavorable (mGCS score < 5).

Data preprocessing

In our data preprocessing approach, variables with missing values exceeding 40% were identified as unreliable and thus removed from the dataset to ensure data integrity. Outliers were then detected using the Interquartile Range (IQR) method, computed as the difference between the 75th and 25th percentiles (Q3 and Q1, respectively). Data points falling outside the range of Q1 - 1.5 * IQR or Q3 + 1.5 * IQR were flagged as outliers and subsequently eliminated based on statistical conventions. Finally, the multiple imputation method was employed for the imputation of missing numerical values. Renowned for its robustness

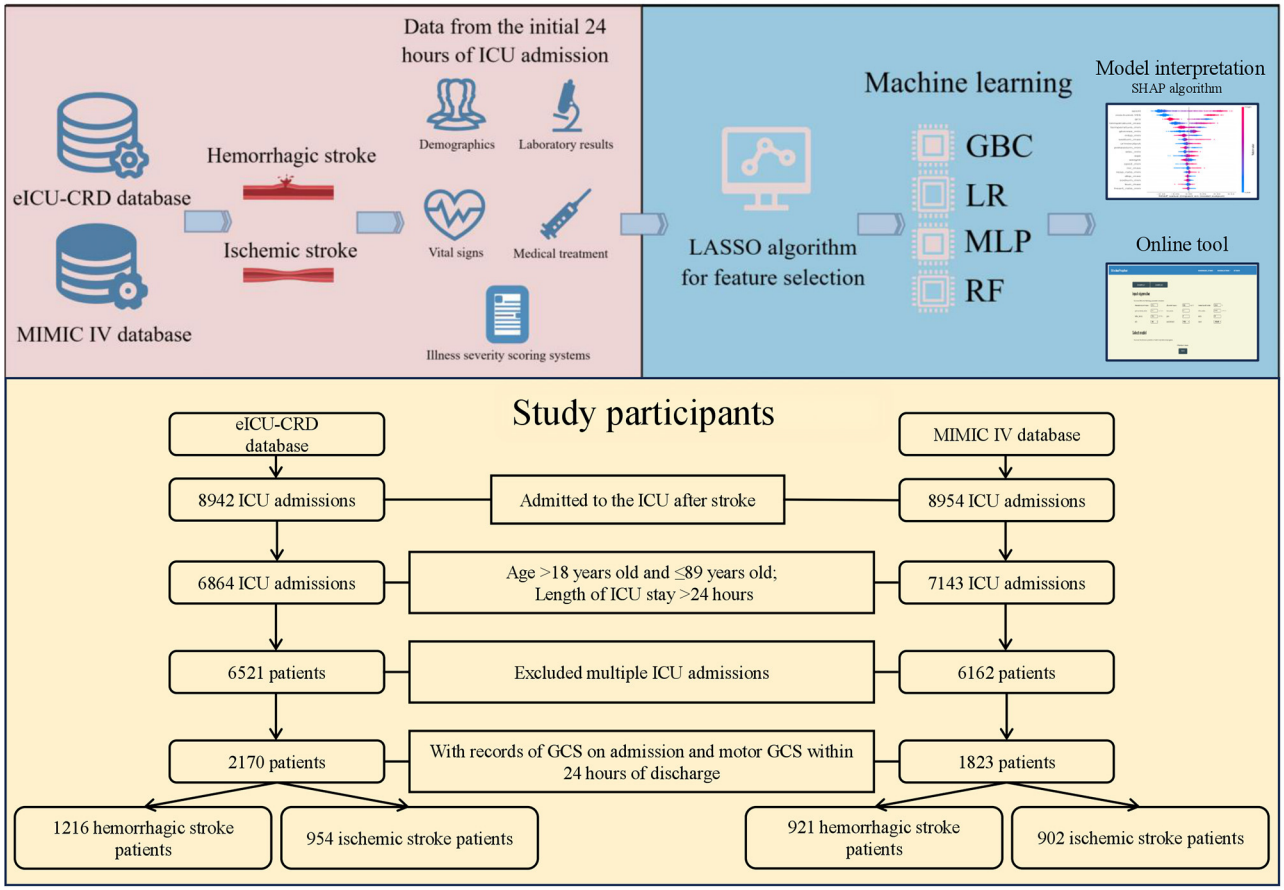


FIGURE 1
Workflow of the study and flow diagram of study participants inclusion.

and capacity to handle intricate datasets, this algorithm efficiently imputes missing values while preserving the inherent data structure.

Model development

In order to prevent overfitting and simplify the model, we simplified the feature data for each outcome, thereby enabling the models to identify underlying patterns in the data and enhance their generalization capability. To achieve this goal, we utilized the least absolute shrinkage and selection operator (LASSO), a machine learning algorithm, and selected the optimal regularization coefficient lambda through a cross-validation process. Specifically, we opted 10-fold for cross-validation to determine the regularization parameter λ (penalty parameter) in the LASSO algorithm. This parameter facilitates variable selection and shrinkage, enabling the compression of some non-essential variables to zero once λ surpasses a certain threshold, thereby excluding them from the model. Through the computation of performance metrics, such as mean squared error, across various λ values during the cross-validation process, we identified the λ value corresponding to the minimum mean squared error as the final regularization parameter. At this λ value, the non-zero coefficients denote the selected significant features. Our study involved four machine learning models, namely Gradient Boosting Classifier (GBC), Logistic Regression (LR), Multi-Layer Perceptron (MLP), and Random Forest (RF). They are all classification models in supervised learning, with GBC being an ensemble learning model, MLP being a deep learning model, while LR and RF are traditional machine learning models. To optimize model performance, we performed hyperparameter tuning with pre-set hyperparameters (Supplementary Tables 1, 2). The hyperparameter tuning was carried out using 10-fold cross-validation during the training set loop.

External validation

We conducted external validation using the MIMIC IV dataset, ensuring consistency in patient inclusion criteria and data processing methods with those described for the eICU patient data. Additionally, we ensured that the clinical indicators analyzed in the MIMIC database maintained consistent units with those in the eICU patient data to ensure the accuracy of validation.

TABLE 1 The international classification of diseases codes of the patients included in the study.

Stroke subtype	ICD code	ICD version	Description
Hemorrhagic stroke	430	9	Subarachnoid hemorrhage
	431	9	Intracerebral hemorrhage
	432	9	Other and unspecified intracranial hemorrhage
	I60	10	Subarachnoid hemorrhage
	I61	10	Intracerebral hemorrhage
	I62	10	Other non-traumatic intracranial hemorrhage
Ischemic stroke	433	9	Occlusion and stenosis of precerebral arteries
	434	9	Occlusion of cerebral arteries
	I63	10	Cerebral infarction
	I65	10	Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
	I66	10	Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction

TABLE 2 Clinical features overview.

Categories	Features
Demographics	Age, gender, race, weight, height, and BMI
Vital signs	Heart rate, systolic blood pressure, diastolic blood pressure, mean blood pressure, temperature, respiratory rate, oxygen saturation, and urineoutput
Laboratory results	Anion gap, bicarbonate, creatinine, chloride, glucose, hematocrit, hemoglobin, lactate, platelet, potassium, ptt, inr, pt, sodium, bun, wbc, and calcium
Medical treatment	Mechanical ventilation, renal replacement therapy, and vasopressor
Illness severity scoring systems	Charlson comorbidity index, Glasgow Coma Scale, Acute Physiology Score III, and Sequential Organ Failure Assessment

BMI, body mass index; PTT, partial thromboplastin time; INR, international normalized ratio; PT, prothrombin time; BUN, blood urea nitrogen; WBC, white blood cell.

Statistical analysis

We evaluated the predictive performance of our model by measuring several common performance metrics, including accuracy, positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, F-measure (F1), and area under the curve (AUC). To determine statistical significance, we used a threshold of $P < 0.05$ and applied Two-tailed Student's t-tests or Mann-Whitney U-tests for continuous variables, as well as Chi-squared or Fisher's exact tests for categorical variables.

Results

Participants

A total of 1,216 hemorrhagic stroke patients and 954 ischemic stroke patients from the eICU-CRD dataset, as well as 921 hemorrhagic stroke patients and 902 ischemic stroke patients from the MIMIC IV dataset, were included in this study. The comparison of baseline features is presented in Table 3, while Supplementary Table 3 provides a summary of specific information for all patients.

In cohort extracted from the eICU-CRD, 33.47% ($n = 407$) of hemorrhagic stroke patients had unfavorable neurological outcome at discharge, while 25.47% ($n = 243$) of ischemic stroke patients had unfavorable neurological outcome at discharge. The proportion of unfavorable neurological outcome for hemorrhagic and ischemic stroke patients from MIMIC IV was 50.71% ($n = 467$) and 40.13% ($n = 362$), respectively. Supplementary Table 4 showed the specific information on hemorrhagic and ischemic stroke patients across training and testing sets.

Feature selection

The results of feature selection based on the LASSO algorithm was shown in Figures 2A–D. The optimal regularization coefficient lambda for each clinical outcome was selected through a cross-validation process. In hemorrhagic cohort, the optimal lambda value for predicting neurological outcome was 0.01023367 and 0.01618563 in ischemic cohort. Figure 2E showed a Venn diagram of the features selected to predict neurological outcome in hemorrhagic and ischemic stroke. Upon application of the LASSO algorithm, a total of 26 and 23 features were discerned to be associated with neurological outcomes in patients diagnosed with hemorrhagic stroke and ischemic stroke, respectively. Encouragingly, it was observed that 13 features exhibited shared significance across both stroke types, emphasizing potential converging mechanisms influencing neurological outcomes.

Model performance

Four models, GBC, LR, MLP, and RF were generated to predict neurological outcome in the extracted cohort using the

TABLE 3 Baseline characteristics of included patients.

	Hemorrhagic stroke			Ischemic stroke		
	eICU-CRD ($n = 1,216$)	MIMIC IV ($n = 921$)	*P*-value	eICU-CRD ($n = 954$)	MIMIC IV ($n = 902$)	*P*-value
Demographic						
Age	64.47 (15.53)	65.63 (15.04)	0.0839	66.38 (14.16)	67.75 (14.02)	0.037
Gender			0.00104			0.00682
Female	656 (53.95 %)	430 (46.69 %)		502 (52.62 %)	417 (46.23 %)	
Male	560 (46.05 %)	491 (53.31 %)		452 (47.38 %)	485 (53.77 %)	
Race			<0.001			<0.001
Asian	17 (1.40 %)	46 (4.99 %)		6 (0.63 %)	23 (2.55 %)	
Black	212 (17.43 %)	90 (9.77 %)		110 (11.53 %)	89 (9.87 %)	
Hispanic	23 (1.89 %)	40 (4.34 %)		11 (1.15 %)	25 (2.77 %)	
Other/unknown	74 (6.09 %)	233 (25.30 %)		35 (3.67 %)	207 (22.95 %)	
White	890 (73.19 %)	512 (55.59 %)		792 (83.02 %)	558 (61.86 %)	
Weight	82.61 (22.55)	80.01 (22.49)	0.00896	83.88 (23.25)	82.31 (23.30)	0.149
Severity scores on admission						
Charlson comorbidity index	3.37 (2.48)	5.29 (2.77)	<0.001	3.76 (2.35)	6.11 (2.90)	<0.001
GCS	11.43 (4.37)	12.40 (3.84)	<0.001	12.34 (3.61)	13.03 (3.36)	<0.001
APSIII	46.29 (26.13)	43.49 (23.86)	0.0113	43.15 (24.15)	44.07 (23.93)	0.422
SOFA	3.57 (2.76)	4.21 (3.65)	<0.001	3.28 (2.88)	4.30 (3.81)	<0.001
First day treatment						
Vasopressor	62 (5.10 %)	186 (20.20 %)	<0.001	73 (7.65 %)	197 (21.84 %)	<0.001
Renal replacement therapy	13 (1.07 %)	24 (2.61 %)	0.0114	13 (1.36 %)	23 (2.55 %)	0.092
Mechanical ventilation	445 (36.60 %)	421 (45.71 %)	<0.001	277 (29.04 %)	303 (33.59 %)	0.191
Hospital length of stay, day	8.78 (8.74)	8.57 (9.72)	0.622	7.50 (11.59)	9.30 (11.46)	<0.001
ICU length of stay, day	5.86 (6.39)	7.61 (8.06)	<0.001	4.44 (5.72)	7.90 (9.86)	<0.001
Neurological outcome						
Favorable	809 (66.53 %)	454 (49.29 %)	<0.001	711 (74.53 %)	540 (59.87 %)	<0.001
Unfavorable	407 (33.47 %)	467 (50.71 %)		243 (25.47 %)	362 (40.13 %)	

Data are n (%) or mean (SD).
GCS, Glasgow Coma Scale; APSIII, Acute Physiology Score III; SOFA, Sequential Organ Failure Assessment.

selected features. Acute physiology and chronic health evaluation IV (APACHE IV), a wildly used method for evaluating critically ill patients, was assessed to compared with the generated models in testing set. A set of detailed performance metrics for various machine learning models was presented in Table 4. Figures 3A, C depicted the predictive performances of the four models and APACHE reference in terms of AUC curve and decision curve analysis (DCA) curve. Among the four models, LR model showed the highest accuracy (0.83), PPV (0.734), specificity (0.86), F1 (0.752) and AUC (0.887). The performance of the optimal model was improved compared to APACHE reference. According to the DCA curves of the four predictive models, the net benefit for LR model was larger over the range of the other models.

In the same way, the four models were generated to predict neurological outcome in ischemic stroke cohort. Figures 3B, D exhibited the discrimination performance of these models via AUC and DCA curves in the testing set. The predictive performance of each model was presented in Table 5. Of the four models, RF model had the best predictive performance (AUC = 0.867). Besides, RF model had the highest F1 (0.66). In addition, GBC model had the highest NPV (0.973) and sensitivity (0.945) and LR model had the highest accuracy (0.794), PPV (0.57) and specificity (0.799).

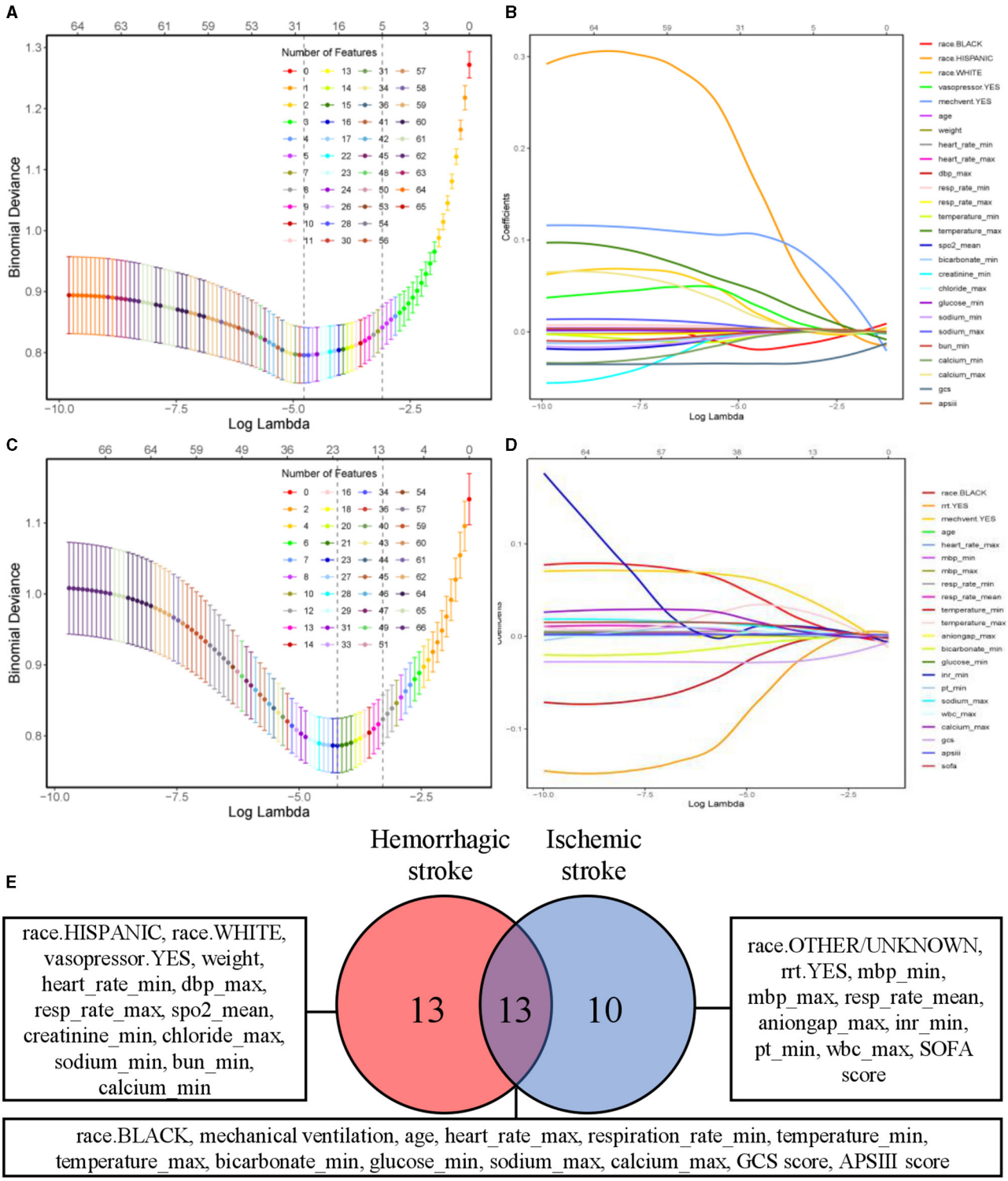


FIGURE 2
Feature selection using LASSO algorithm. Tuning parameter λ selection by 10-fold cross-validation with minimum criteria. The binomial deviance is plotted on the y-axis against the logarithm of λ on the x-axis. Vertical dotted lines are drawn at the optimal value of λ, which corresponds to the point where the model achieves the best fit to the data. **(A, C)** Depicting partial likelihood deviance of the LASSO regression for neurological outcome of patients with hemorrhagic and ischemic stroke, respectively. **(B, D)** Coefficient profile of the clinical features associated with neurological outcome of patients with hemorrhagic stroke and ischemic stroke. **(E)** Venn diagram of selected features associated with neurological outcome in hemorrhagic and ischemic stroke. LASSO, least absolute shrinkage and selection operator.

TABLE 4 Model performance summary of all models in hemorrhagic stroke test cohort.

Model	Accuracy	PPV	NPV	Sensitivity	Specificity	F1	AUC
GBC	0.814	0.68	0.907	0.836	0.802	0.750	0.865
LR	0.830	0.734	0.882	0.770	0.860	0.752	0.887
MLP	0.781	0.635	0.890	0.811	0.765	0.712	0.855
RF	0.789	0.645	0.895	0.820	0.774	0.722	0.868
APACHE reference	0.817	0.716	0.869	0.735	0.858	0.725	0.868

GBC, gradient boosting classifier; LR, logistic regression; MLP, multi-layer perceptron; RF, random forest; APACHE, acute physiology and chronic health evaluation; PPV, positive predictive value; NPV, negative predictive value; F1, F-measure; AUC, area under curve.

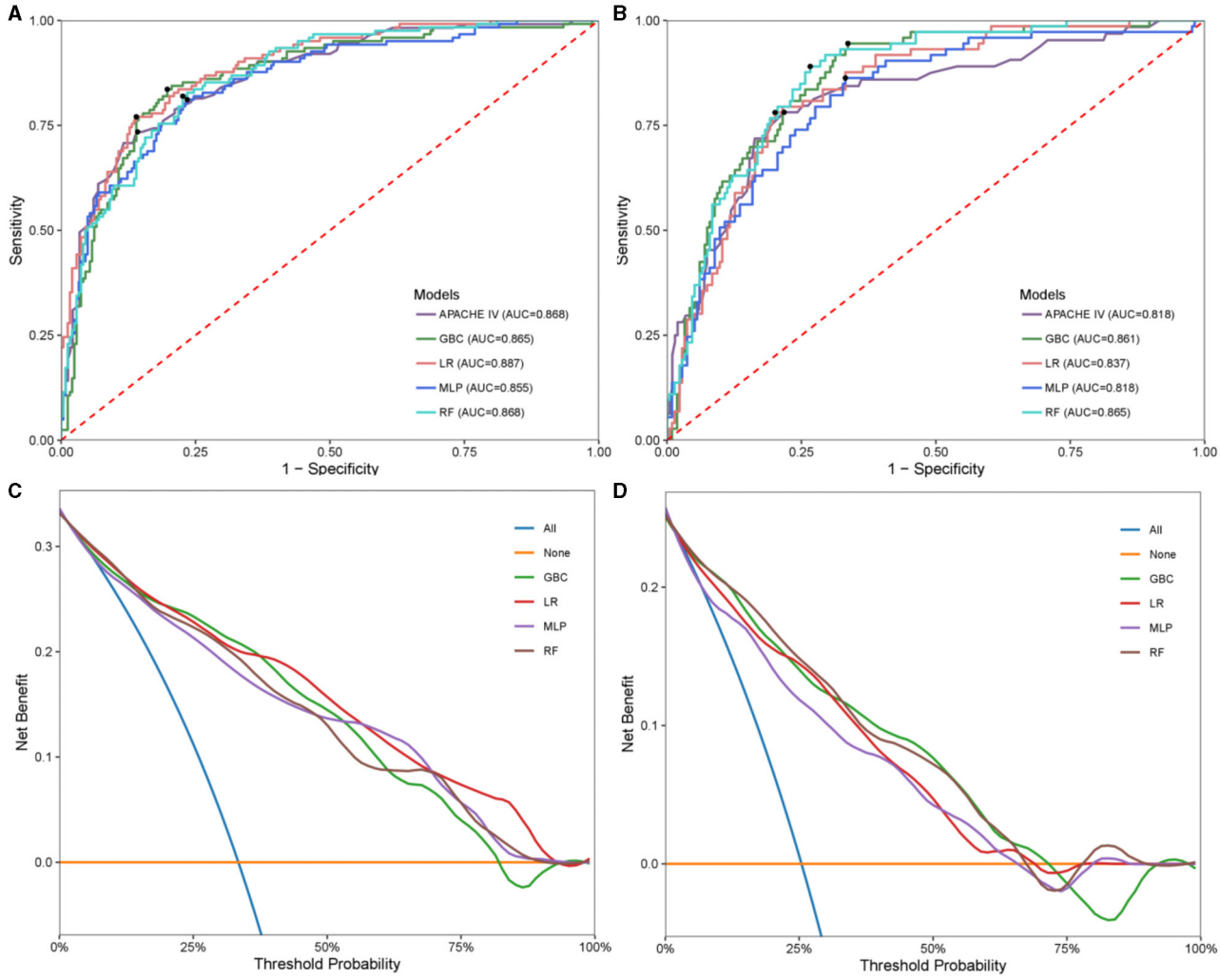


FIGURE 3
Model performance in hemorrhagic stroke test cohort and ischemic stroke test cohort. Receiver operating characteristic (ROC) analysis of GBC, LR, MLP, RF models, and APACHE reference. **(A)** Hemorrhagic stroke; **(B)** Ischemic stroke. Decision curve analysis (DCA) curves of four machine learning models. **(C)** Hemorrhagic stroke; **(D)** Ischemic stroke. GBC, gradient boosting classifier; LR, logistic regression; MLP, multi-layer perceptron; RF, random forest; APACHE, acute physiology and chronic health evaluation.

Model interpretation

In order to comprehensively elucidate the effect of various clinical features on the neurological outcome of stroke patients, we employed the SHAP algorithm to determine their overall positive or negative impact on the optimal model output. As shown in Figure 4A, GCS score ranked the first in importance among the features for predicting neurological outcome in hemorrhagic stroke cohort, followed by APS III score, age, glucose_min, and sodium_max. Figure 4B showed that GCS score had the most potent predictive power in predicting neurological outcome in ischemic stroke patients, followed by APS III score, SOFA

TABLE 5 Model performance summary of all models in ischemic stroke test cohort.

Model	Accuracy	PPV	NPV	Sensitivity	Specificity	F1	AUC
GBC	0.735	0.489	0.973	0.945	0.664	0.645	0.861
LR	0.794	0.570	0.914	0.781	0.799	0.659	0.837
MLP	0.718	0.470	0.935	0.863	0.668	0.609	0.818
RF	0.763	0.520	0.956	0.904	0.715	0.660	0.867
APACHE reference	0.782	0.532	0.919	0.781	0.782	0.633	0.818

GBC, gradient boosting classifier; LR, logistic regression; MLP, multi-layer perceptron; RF, random forest; PPV, positive predictive value; NPV, negative predictive value; F1, F-measure; AUC, area under curve.

score, mechanical ventilation and temperature_max. To offer a comprehensive overview of feature importance ranking in optimal model construction, Supplementary Figure 1 also provided the summaries of the feature importance ranking for predicting neurological outcome in hemorrhagic and ischemic stroke.

To facilitate a more intuitive understanding of how alterations in individual clinical features impact the model's output, we included SHAP dependence plots that depict the top 10 contributing features for each model in Figure 5, SHAP dependence plots of the remaining features were shown in Supplementary Figures 2, 3. In the case of hemorrhagic stroke, as illustrated in Figures 5A–J, patients with lower GCS score (<11), sodium_min (<137 mmol/L), SpO_2_mean (<97%) and lighter weight (<80 kg) or higher APS III score (>50), glucose_min (>125 mg/dL), sodium_max (>142 mmol/L), heart rate_max (>100 beats/minute), temperature_max (>37.5°C), and order age (>65 years) are more likely to be predicted as having unfavorable neurological outcome. In the ischemic stroke group, the effects of GCS score, APS III score, temperature_max, heart rate_max, glucose_min, and sodium_max on the model's predictions aligned with those observed in the hemorrhagic stroke group. Additionally, ischemic stroke patients with higher SOFA score (>3) and wbc_max ($>10^{9/L}$), lower bicarbonate_min (<23 mEq/L) or treated with mechanical ventilation are more prone to be predicted as having unfavorable neurological outcome (Figures 5K–T).

External validation

Our predictive analyses of the MIMIC IV validation cohorts demonstrated a consistent pattern with that observed in the eICU-CRD cohorts (Supplementary Figure 4). In the validation set, the LR model demonstrated an AUC of 0.836 for predicting neurological outcome in hemorrhagic stroke patients, which represented a decrease of 0.051 compared to the testing set. For the ischemic stroke validation set, the AUC of the RF model was 0.856 for neurological outcome, demonstrating a reduction of 0.011. A comprehensive overview of external validation results is in Supplementary Tables 5, 6.

Online tool for prediction

Based on the optimal model for predicting neurological outcome in hemorrhagic and ischemic stroke, along with the relevant clinical variables that it encompasses, we devised an online prediction tool (Figure 6). By selecting the stroke type and entering the relevant clinical data, the user can obtain an automatically generated ID for outcome query. After entering the above ID into the query interface, the user can obtain the neurological outcome prediction result. This prediction tool is accessible at: http://www.strokeprophet.cn.

Discussion

This retrospective analysis delved into the medical records of stroke patients, utilizing data from the eICU-CRD multicenter database, and effectively validated the findings using the MIMIC IV database. By employing multiple algorithmic techniques and various machine learning models, we successfully identified the features of clinical indicators within the first 24 h of admission that are highly correlated with the neurofunctional state at discharge. Importantly, we integrated the SHAP algorithm and existing literature to provide comprehensive interpretability and inference for these factors, which holds significant value in guiding future prospective research and supporting clinical decision-making in stroke care. Most importantly, the results of this study have been translated into a practical online tool that enables precise prediction of neurologic prognosis in hemorrhagic and ischemic stroke patients using features from the first 24 h in the ICU.

Categorically, stroke can be divided into two primary subtypes: ischemic stroke, accounting for 87% of all cases, and hemorrhagic stroke, comprising intracerebral hemorrhage and subarachnoid hemorrhage, which collectively make up 10 and 3% of stroke occurrences, respectively. Prompt emergency interventions aimed at restoring blood flow prove crucial in improving patient prognosis in the context of ischemic stroke. Conversely, effective management of hemorrhagic stroke necessitates surgical hemostasis and the control of intracranial pressure (14). Existing studies primarily consist of single-center, retrospective observational research aimed at understanding the differences in neurofunctional influencing factors between these two subtypes of stroke (14). However, there is a lack of reported research investigating the neurofunctional prediction and exploration of risk factors specifically in these two stroke patient populations. Our study significantly contributes to filling this gap in knowledge.

Our comprehensive study reveals that within 24 h of hospital admission, ventilator use, age, temperature, GCS score, APS III score, heart rate, blood sodium levels, blood calcium levels, respiratory rate, blood bicarbonate levels, blood glucose

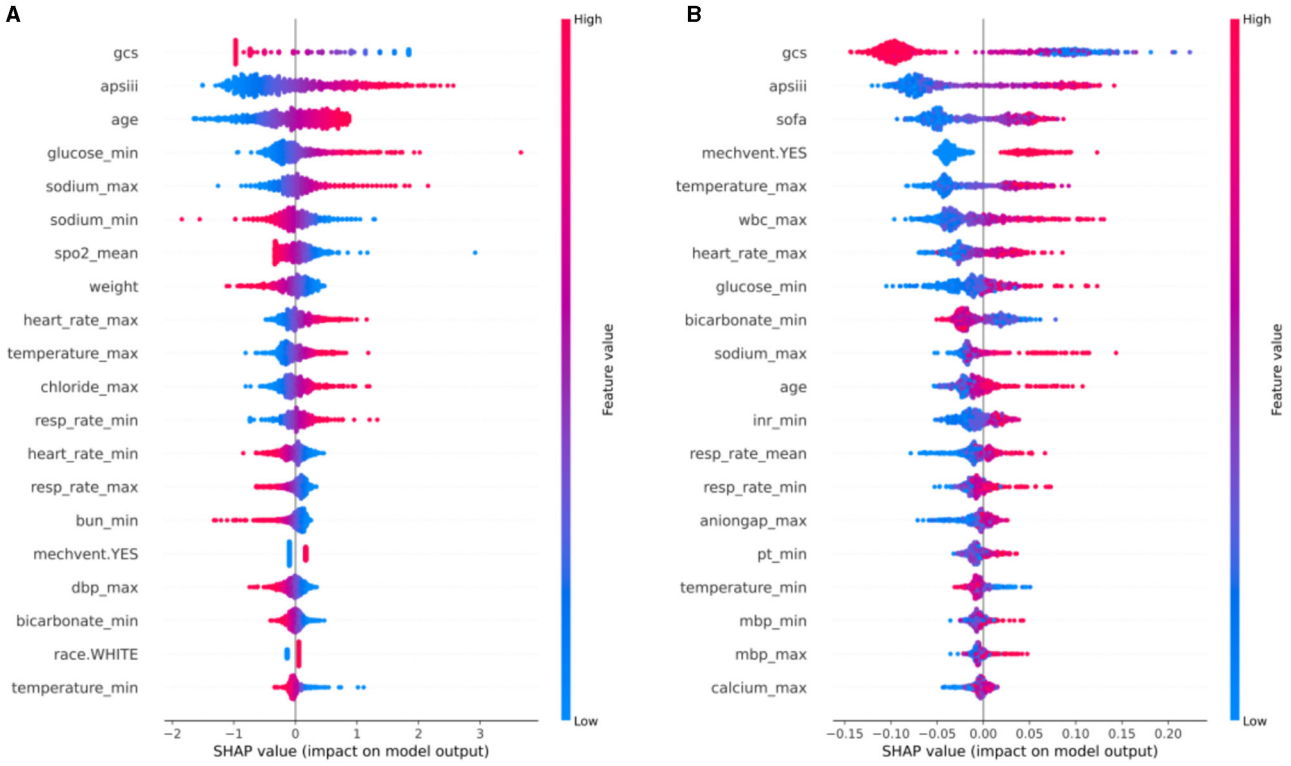


FIGURE 4
SHAP analysis of the optimal model for neurological outcome in hemorrhagic stroke cohort **(A)** and ischemic stroke cohort **(B)**. SHAP, Shapley additive explanations.

levels, and ethnicity all demonstrate noteworthy significance in influencing the prognosis of neurological function for both stroke types. It is imperative to carefully scrutinize and monitor these factors in patients diagnosed with either hemorrhagic or ischemic stroke to facilitate a more accurate assessment of their neurological prognosis.

Recent discoveries indicate that the requirement for mechanical ventilation among stroke patients may exhibit a stronger correlation with the site of brain injury rather than the stroke subtype itself. Consequently, optimizing ventilation strategies assumes a crucial role in enhancing the prognosis of patients within the hemorrhagic stroke cohort (15). In stark contrast, our investigation unveils an intriguing observation: aside from its influence on hemorrhagic stroke outcomes, mechanical ventilator employment also exerts a significant detrimental impact on the prognosis of neurological function in patients suffering from ischemic stroke. Furthermore, our findings identify respiratory rate as a contributing factor affecting the neurological prognosis in both stroke subtypes. These observations may be germane to the development of complicating conditions such as ventilator-associated pneumonia and acute respiratory distress syndrome (16). Therefore, when treating stroke patients, individual circumstances should be carefully considered, mechanical ventilation therapy should be optimized.

Several studies have yielded compelling evidence regarding the influence of hypernatremia and hypernatremia on the neurological prognosis of stroke patients (17, 18). Notably, we have identified a novel association between blood bicarbonate ion levels and the neurological prognosis in both hemorrhagic and ischemic stroke patients, a correlation that has not been previously reported in the literature. The presence of acute hypernatremia and underlying diabetes, recognized risk factors for cerebrovascular disease, further exacerbate the neurological prognosis in stroke patients (19). Additionally, heightened heart rate and elevated body temperature emerge as crucial factors contributing to an unfavorable prognosis among this patient population (20, 21). Our investigation also reveals that advanced age significantly impairs neurological recovery, with older stroke patients exhibiting diminished prognosis relative to their younger counterparts (22). Intriguingly, the observations in elderly mice suggest an enhanced propensity for neutrophil plugging in the ischemic brain microcirculation post-stroke, resulting in compromised blood flow and worsened prognosis (23). Furthermore, dysfunctionality within microglia in aged mouse brains may contribute to the deteriorating neurological prognosis following stroke (24). Similarly, Peng et al.'s report on stroke cohorts from 1990 to 2019 also indicates a relative inadequacy in healthcare for stroke patients across all age groups, underscoring the need for further investigation into the impact of age on mortality trends (25). Disparities in neurological prognosis between racial groups have been noted, with Black patients exhibiting poorer outcomes compared to White patients, potentially linked to stroke subtype (26). Furthermore, Huang et al.'s study also found that race is one of the top 11 most important

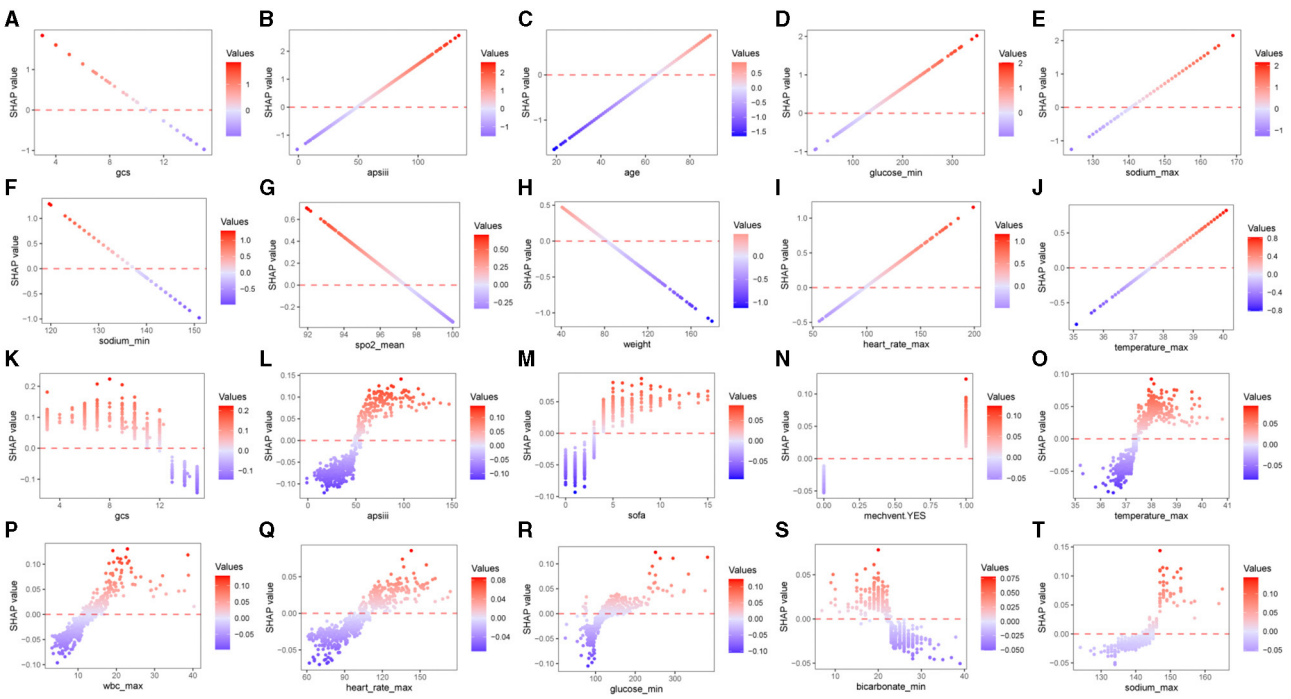


FIGURE 5
SHAP dependence plots of the top 10 features of predicting neurological outcome in hemorrhagic stroke cohort **(A–J)** and ischemic stroke cohort **(K–T)**. SHAP, Shapley additive explanations; GCS, Glasgow Coma Scale; APS III, Acute Physiology Score III; SpO_2, Oxygen saturation; SOFA, Sequential Organ Failure Assessment; WBC, white blood cell.

features for predicting 28-day all-cause in-hospital mortality among hypertensive ischemic or hemorrhagic stroke patients (27). Our study corroborates these findings, highlighting that White patients and Hispanic patients experience inferior neurological prognosis in hemorrhagic stroke, while other ethnic groups fare worse in ischemic stroke. Remarkably, Black patients exhibit poorer prognoses across both stroke subtypes (28). Nevertheless, comprehensive investigations are warranted to validate and elucidate these observations.

Scoring systems play a pivotal role in evaluating the condition and predicting the prognosis of patients, thereby guiding treatment strategies and facilitating informed decision-making. Among these scoring systems, the GCS has widespread application in stroke patients, individuals undergoing open-heart surgery, and those with varying degrees of coma arising from diverse etiologies (29). Our investigation highlights a significant correlation between GCS score and APS III scores and the neurological outcome of patients afflicted by both hemorrhagic and ischemic stroke. Notably, hemorrhagic stroke patients exhibited diminished neurological prognosis when their GCS score was $\leq$11, whereas ischemic stroke patients experienced compromised functional prognosis when their GCS score fell below $\leq$12. However, the underlying mechanisms contributing to these findings necessitate further inquiry. Furthermore, our study reveals that the SOFA score exerts a more pronounced impact on the neurological prognosis of individuals with ischemic stroke relative to the aforementioned scoring metrics (30).

Electrolyte imbalance manifests earlier in individuals suffering from ischemic stroke and holds the potential to serve as a prognostic indicator for neurological outcomes among stroke patients (31). In light of our investigation, it has been established that excessive serum chloride ion concentration (>110 mmol/L) exerts a more pronounced influence on neurological function within the hemorrhagic stroke patient cohort, thereby emphasizing the role of averting hyperchloremia in enhancing neurological outcomes for individuals with hemorrhagic stroke (32). Remarkably, our novel observation reveals that heightened anion gap exhibits a greater impact on the neurological prognosis of ischemic stroke patients compared to ischemic individuals, an unprecedented discovery in current literature. Clinical practitioners should therefore dedicate substantial attention to blood gas analysis results, ensuring prompt correction of electrolyte imbalances while avoiding overcorrection.

Studies in the field of stroke hemodynamics have predominantly concentrated on ischemic stroke due to its higher incidence rate compared to hemorrhagic stroke. Both hypertension and hypotension have detrimental effects on acute ischemic stroke. Generally, antihypertensive therapy is recommended when systolic blood pressure (SBP) exceeds 220 or 180 mmHg in patients eligible for thrombolytic therapy. In the case of acute hemorrhagic stroke, SBP exceeding 140 mmHg is associated with an unfavorable neurological prognosis (33). Notably, our investigation reveals that when monitoring circulatory dynamics in stroke patients, it is crucial to adopt distinct approaches depending on whether the patient has a hemorrhagic or ischemic stroke. For patients with hemorrhagic stroke, the prognostic significance of maximum diastolic blood pressure on neurological function is of greater importance. Conversely, in ischemic patients, both the highest and

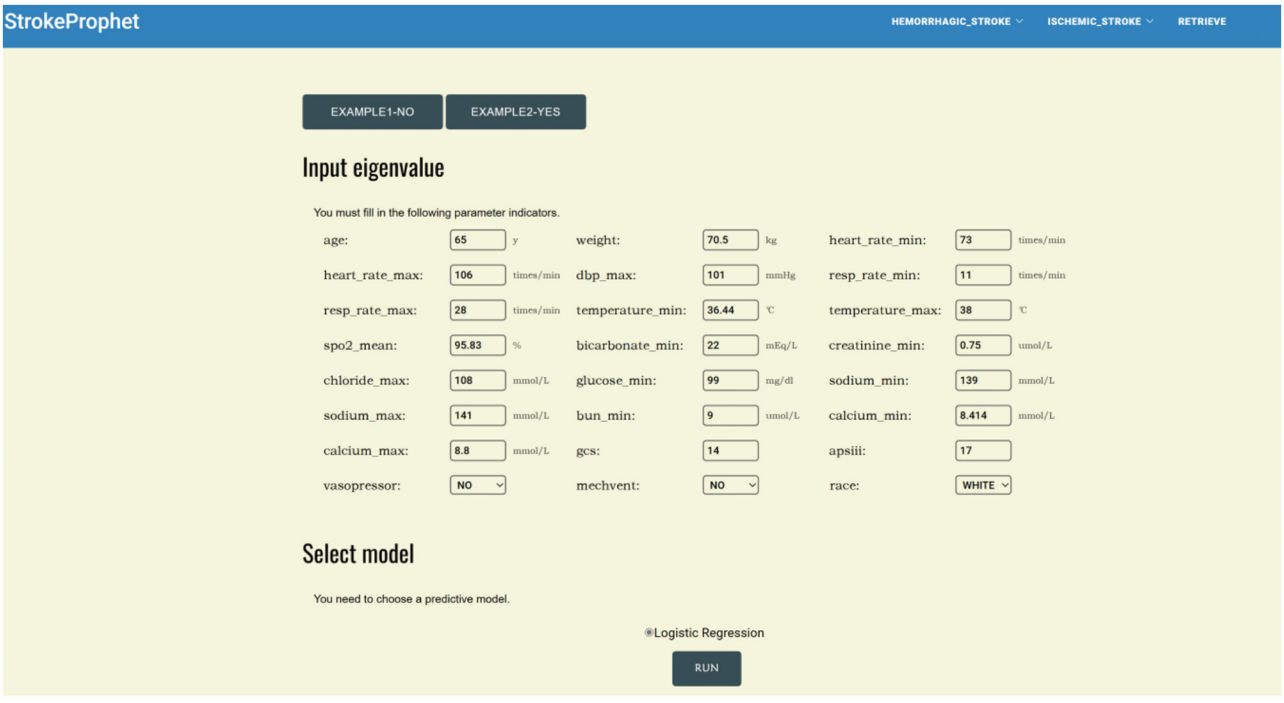


FIGURE 6
An example of online tool for prediction.

lowest mean blood pressure levels exert an impact on prognosis. Furthermore, the administration of vasopressors escalates the risk of poor neurological prognosis among individuals with hemorrhagic stroke.

In conjunction with the aforementioned findings, it is imperative to consider several additional factors that warrant separate attention and pertain to patients afflicted with different stroke types. Among individuals with hemorrhagic stroke, maintaining an optimal level of SpO_2 proves advantageous for prognosis. Notably, accumulating evidence suggests that oxygen therapy is not devoid of risks and should be withheld when SpO_2 exceeds 90%. Moreover, in patients at risk of hypercapnia, the threshold for oxygen therapy ought to be even lower, specifically $\leq$ 88%. Once oxygen therapy has been initiated, diligent monitoring of the patient's oxygen saturation and inhaled oxygen concentration becomes paramount in order to maintain SpO_2 within the targeted range (93–96%) and avert the detrimental consequences of hypercapnia (34). Furthermore, it is worth noting that hemorrhagic stroke patients with lower body weight exhibit an inferior neurological prognosis (35), while diminished levels of blood urea nitrogen are likewise associated with an unfavorable neurological outcome (36).

In the context of individuals suffering from ischemic stroke, heightened emphasis should be placed on monitoring the patient's blood leucocyte count as a means to mitigate infection risk and enhance neurological prognosis (37). Moreover, it is crucial to duly acknowledge the substantial influence exerted by the international normalized ratio (INR) and plasma prothrombin time on ischemic stroke patients, especially following thrombolytic therapy, warranting careful attention (38, 39). These discoveries offer valuable insights into the management of stroke patients necessitating intensive care unit treatment and furnish guidance for improved resource allocation in clinical decision-making.

The final point we wish to emphasize is that machine learning algorithms excel at constructing complex models and making informed decisions when provided with ample and relevant data. In our study, the choice of machine learning models was carefully guided by their unique strengths and documented effectiveness in similar tasks. The GBC model was selected for its exceptional performance in managing high-dimensional

data and intricate feature relationships. By leveraging ensemble learning techniques, GBC adeptly captures non-linear associations within the data, making it well-suited for addressing complex classification challenges (40). The RF model was chosen owing to its resilience and efficiency in handling high-dimensional datasets with numerous features (41). Notably, Elsaid et al.'s study on hemorrhagic transformation prediction found that both RFC and GBC models, capable of capturing non-linear interactions among predictor variables, yielded the best predictive performance (AUC: 0.91, 95% CI: 0.85–0.95; AUC: 0.91, 95% CI: 0.86–0.95, respectively) (42). For binary classification tasks, the simplicity, interpretability, and effectiveness of LR made it a prudent choice. Su et al.'s investigation on post-stroke cognitive impairment (PSCI) underscored LR's efficacy in discerning the varying impacts of different factors on cognitive impairment (43). The MLP model was selected for its ability to handle complex non-linear relationships, as demonstrated in Zhou et al.'s study on dementia cognitive footprint recognition (44). MLP emerged as the top-performing model, showcasing satisfactory performance in dementia identification. In our study, the LR model demonstrated the highest AUC of 0.887 in the hemorrhagic stroke cohort, whereas the RF model exhibited superior performance in the ischemic stroke cohort, with an AUC of 0.867. Considering the insightful observations derived from SHAP regarding the importance of variables in predicting each stroke subtype, we analyze and speculate that the primary reason for the differences in optimal models lies in the interactions among predictive factors. Specifically, in hemorrhagic stroke prediction, the most crucial variables predominantly consisted of continuous variables such as age, weight, and glucose. This suggests that LR, with its linear decision boundary, may be better suited to capture the linear relationships among these predictors, thereby achieving higher predictive performance. Conversely, in ischemic stroke prediction, notable variables included the use of mechanical ventilation (mechvent). This intriguing finding suggests that RF, with its capacity to capture complex non-linear relationships, may excel in identifying intricate patterns involving categorical variables like mechvent, which LR may overlook due to its linear nature. These findings lead us to hypothesize that LR's superior performance in hemorrhagic stroke prediction may be attributed to the predominantly linear relationships among predictors, while RF's effectiveness in predicting ischemic stroke-induced damage may stem from its ability to capture non-linear associations. However, further research and validation are warranted to confirm these hypotheses and gain deeper insights into the underlying mechanisms driving model performance discrepancies between stroke subtypes.

The present study has several limitations that warrant consideration. Firstly, our analysis was retrospective in nature, which resulted in the exclusion of certain variables with high missing rates but potentially significant predictive value, such as lactate levels. Secondly, our inability to directly evaluate patients necessitated reliance on diagnostic codes to define our patient cohort, raising the possibility of incorrect associations due to misclassifications. Thirdly, our focus on patients who remained in the ICU for over 24 h and had an mGCS score 24 h before discharge resulted in the exclusion of a sizable number of patients, and this may have introduced some bias into our results. In addition, some important features, namely neuroimaging and electrophysiological examinations, were not included in this study due to the limitations of the database.

Conclusions

In summary, we employed advanced machine learning techniques to identify and compare the shared and distinct factors that influence hospital discharge outcomes in both hemorrhagic and ischemic stroke. By elucidating these factors, our research will contribute to the advancement of knowledge in the field of stroke, inform medical decision-making, and guide personalized treatment strategies. Our results also demonstrate that machine learning models outperform single standard scoring systems, with the potential to ultimately improve patient outcomes.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

HW: Writing—original draft, Writing—review & editing. XH: Writing—original draft, Writing—review & editing. YZ: Writing—original draft, Writing—review & editing. GJ: Writing—original draft, Writing—review & editing. RD: Writing—review & editing. MD: Writing—review & editing. LW: Writing—review & editing. HY: Writing—original draft, Writing—review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by Military Logistics Research Projects (Grant No. AHJ21J003) and the Science and Technology Commission of Shanghai Municipality (Grant No. 22xtcx00603).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1385013/full#supplementary-material

References

1. GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* (2019) 18:459–80. doi: 10.1016/S1474-4422(18)30499-X

2. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* (2021) 20:795–820. doi: 10.1016/S1474-4422(21)00252-0

3. Zhang X, Fei N, Zhang X, Wang Q, Fang Z. Machine learning prediction models for postoperative stroke in elderly patients: analyses of the MIMIC database. *Front Aging Neurosci.* (2022) 14:897611. doi: 10.3389/fnagi.2022.897611

4. Joundi R, Smith EE, Yu AYX, Rashid M, Fang J, Kapral MK. Age-specific and sex-specific trends in life-sustaining care after acute stroke. *J Am Heart Assoc.* (2021) 10:e021499. doi: 10.1161/JAHA.121.021499

5. Mayer SA, Copeland D, Bernardini GL, Boden-Albala B, Lennihan L, Kossoff S, et al. Cost and outcome of mechanical ventilation for life-threatening stroke. *Stroke.* (2000) 31:2346–53. doi: 10.1161/01.STR.31.10.2346

6. van Valburg MK, Arbous MS, Georgieva M, Brealey DA, Singer M, Geerts BF. Clinical predictors of survival and functional outcome of stroke patients admitted to critical care. *Crit Care Med.* (2018) 46:1085–92. doi: 10.1097/CCM.0000000000003127

7. Diaz JV, Riviello ED, Papali A, Adhikari NKJ, Ferreira JC. Global critical care: moving forward in resource-limited settings. *Ann Glob Health.* (2019) 85:2413. doi: 10.5334/aogh.2413

8. 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) 50:1263–5. doi: 10.1161/STROKEAHA.118.024293

9. Xu X, Qu J, Zhang Y, Qian X, Chen T, Liu Y. Development and validation of an MRI-radiomics nomogram for the prognosis of pancreatic ductal adenocarcinoma. *Front Oncol.* (2023) 13:1074445. doi: 10.3389/fonc.2023.1074445

10. van Os H, Ramos LA, Hilbert A, van Leeuwen M, van Walderveen MAA, Kruyt ND, et al. Predicting outcome of endovascular treatment for acute ischemic stroke: potential value of machine learning algorithms. *Front Neurol.* (2018) 9:784. doi: 10.3389/fneur.2018.00784

11. Zhang Y, Zhu S, Yuan Z, Li Q, Ding R, Bao X, et al. Risk factors and socio-economic burden in pancreatic ductal adenocarcinoma operation: a machine learning based analysis. *BMC Cancer.* (2020) 20:1161. doi: 10.1186/s12885-020-07626-2

12. Peng S, Huang J, Liu X, Deng J, Sun C, Tang J, et al. Interpretable machine learning for 28-day all-cause in-hospital mortality prediction in critically ill patients with heart failure combined with hypertension: a retrospective cohort study based on medical information mart for intensive care database-IV and eICU databases. *Front Cardiovasc Med.* (2022) 9:994359. doi: 10.3389/fcvm.2022.994359

13. Huang J, Jin W, Duan X, Liu X, Shu T, Fu L, et al. Twenty-eight-day in-hospital mortality prediction for elderly patients with ischemic stroke in the intensive care unit: interpretable machine learning models. *Front Public Health.* (2022) 10:1086339. doi: 10.3389/fpubh.2022.1086339

14. Salvadori E, Papi G, Insalata G, Rinnoci V, Donnini I, Martini M, et al. Comparison between ischemic and hemorrhagic strokes in functional outcome at discharge from an intensive rehabilitation hospital. *Diagnostics.* (2020) 11:38. doi: 10.3390/diagnostics11010038

15. Robba C, Bonatti G, Battaglini D, Rocco PRM, Pelosi P. Mechanical ventilation in patients with acute ischaemic stroke: from pathophysiology to clinical practice. *Crit Care.* (2019) 23:388. doi: 10.1186/s13054-019-2662-8

16. Huang M, Gedansky A, Hassett CE, Price C, Fan TH, Stephens RS, et al. Pathophysiology of brain injury and neurological outcome in acute respiratory distress syndrome: a scoping review of preclinical to clinical studies. *Neurocrit Care.* (2021) 35:518–27. doi: 10.1007/s12028-021-01309-x

17. Yuen KCJ, Sharf V, Smith E, Kim M, Yuen ASM, MacDonald PR. Sodium and water perturbations in patients who had an acute stroke: clinical relevance and management strategies for the neurologist. *Stroke Vasc Neurol.* (2022) 7:258–66. doi: 10.1136/svn-2021-001230

18. Chi C, Patel S, Cheung N. Admission sodium levels and hospital outcomes. *Intern Med J.* (2021) 51:93–8. doi: 10.1111/imj.14777

19. Kim S, Kim JY, Kim ES, Park IR, Ha EY, Chung SM, et al. Early glycaemic variability increases 28-day mortality and prolongs intensive care unit stay in critically ill patients with pneumonia. *Ann Med.* (2022) 54:2736–43. doi: 10.1080/07853890.2022.2128399

20. Siebert J, Gutknecht P, Molisz A, Trzeciak B, Nyka W. Hemodynamic findings in patients with brain stroke. *Archiv Med Sci.* (2012) 8:371–4. doi: 10.5114/aoms.2012.28567

21. Kallmünzer B, Kollmar R. Temperature management in stroke—an unsolved, but important topic. *Cerebrovasc Dis.* (2011) 31:532–43. doi: 10.1159/000324621

22. Roy-O'Reilly M, McCullough L. Age and sex are critical factors in ischemic stroke pathology. *Endocrinology.* (2018) 159:3120–31. doi: 10.1210/en.2018-00465

23. Gullotta G, De Feo D, Friebel E, Semerano A, Scotti GM, Bergamaschi A, et al. Age-induced alterations of granulopoiesis generate atypical neutrophils that aggravate stroke pathology. *Nat Immunol.* (2023) 24:925–40. doi: 10.1038/s41590-023-01505-1

24. Arbaizar-Rovirosa M, Pedragosa J, Lozano JJ, Casal C, Pol A, Gallizioli M, et al. Aged lipid-laden microglia display impaired responses to stroke. *EMBO Mol Med.* (2023) 15:e17175. doi: 10.15252/emmm.202217175

25. Peng S, Liu X, Cao W, Liu Y, Liu Y, Wang W, et al. Global, regional, and national time trends in mortality for stroke, 1990-2019: an age-period-cohort analysis for the global burden of disease 2019 study and implications for stroke prevention. *Int J Cardiol.* (2023) 383:117–31. doi: 10.1016/j.ijcard.2023.05.001

26. Simmonds K, Luo Z, Reeves M. Race/ethnic and stroke subtype differences in poststroke functional recovery after acute rehabilitation. *Arch Phys Med Rehabil.* (2021) 102:1473–81. doi: 10.1016/j.apmr.2021.01.090

27. Huang J, Chen H, Deng J, Liu X, Shu T, Yin C, et al. Interpretable machine learning for predicting 28-day all-cause in-hospital mortality for hypertensive ischemic or hemorrhagic stroke patients in the ICU: a multi-center retrospective cohort study with internal and external cross-validation. *Front Neurol.* (2023) 14:1185447. doi: 10.3389/fneur.2023.1185447

28. Alter M. Black-white differences in stroke frequency: challenges for research. *Neuroepidemiology.* (1994) 13:301–7. doi: 10.1159/000110395

29. Handschu R, Haslbeck M, Hartmann A, Fellgiebel A, Kolominsky-Rabas P, Schneider D, et al. Mortality prediction in critical care for acute stroke: severity of illness-score or coma-scale? *J Neurol.* (2005) 252:1249–54. doi: 10.1007/s00415-005-0853-5

30. Qin W, Zhang X, Yang L, Li Y, Yang S, Li X, et al. Predictive value of the sequential organ failure assessment (SOFA) score for prognosis in patients with severe acute ischemic stroke: a retrospective study. *J Int Med Res.* (2020) 48:300060520950103. doi: 10.1177/0300060520950103

31. Jhou H-J, Chen P-H, Yang L-Y, Chang S-H, Lee C-H. Plasma anion gap and risk of in-hospital mortality in patients with acute ischemic stroke: analysis from the MIMIC-IV database. *J Personal Med.* (2021) 11:1004. doi: 10.3390/jpm11101004

32. Qureshi A, Huang W, Hanley DF, Hsu CY, Martin RH, Malhotra K, et al. Early hyperchloremia is independently associated with death or disability in patients with intracerebral hemorrhage. *Neurocritical Care.* (2022) 37:487–96. doi: 10.1007/s12028-022-01514-2

33. Donovan A, Flexman A, Gelb A. Blood pressure management in stroke. *Curr Opin Anaesthesiol.* (2012) 25:516–22. doi: 10.1097/ACO.0b013e32835721a5

34. Chu D, Kim LH, Young PJ, Zamiri N, Almenawer SA, Jaeschke R, et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet.* (2018) 391:1693–705. doi: 10.1016/S0140-6736(18)30479-3

35. Shiozawa M, Kaneko H, Itoh H, Morita K, Okada A, Matsuoka S, et al. Association of body mass index with ischemic and hemorrhagic stroke. *Nutrients.* (2021) 13:72343. doi: 10.3390/nu13072343

36. Peng R, Liu K, Li W, Yuan Y, Niu R, Zhou L, et al. Blood urea nitrogen, blood urea nitrogen to creatinine ratio and incident stroke: the Dongfeng-Tongji cohort. *Atherosclerosis.* (2021) 333:1–8. doi: 10.1016/j.atherosclerosis.2021.08.011

37. Hu Z-B, Zhong Q-Q, Lu Z-X, Zhu F. Association of platelet-to-white blood cell ratio and platelet-to-neutrophil ratio with the risk of fatal stroke occurrence in middle-aged to older Chinese. *BMC Geriatr.* (2022) 22:430. doi: 10.1186/s12877-022-03134-z

38. Odén A, Fahlén M, Hart R. Optimal INR for prevention of stroke and death in atrial fibrillation: a critical appraisal. *Thromb Res.* (2006) 117:493–9. doi: 10.1016/j.thromres.2004.11.025

39. Huang X, Moreton FC, Kalladka D, Cheripelli BK, MacIsaac R, Tait RC, et al. Coagulation and fibrinolytic activity of tenecteplase and alteplase in acute ischemic stroke. *Stroke.* (2015) 46:3543–6. doi: 10.1161/STROKEAHA.115.011290

40. Ruksakulpiwat S, Thongking W, Zhou W, Benjasirisan C, Phianhasin L, Schiltz NK, et al. Machine learning-based patient classification system for adults with stroke: a systematic review. *Chronic Illn.* (2023) 19:26–39. doi: 10.1177/17423953211067435

41. Hu J, Szymczak S. A review on longitudinal data analysis with random forest. *Brief Bioinform.* (2023) 24:bbad002. doi: 10.1093/bib/bbad002

42. Elsaid AF, Fahmi RM, Shehta N, Ramadan BM. Machine learning approach for hemorrhagic transformation prediction: capturing predictors' interaction. *Front Neurol.* (2022) 13:951401. doi: 10.3389/fneur.2022.951401

43. Su W, Li H, Dang H, Han K, Liu J, Liu T, et al. Predictors of cognitive functions after stroke assessed using the wechsler adult intelligence scale: a retrospective study. *J Alzheimers Dis.* (2024) 98:109–17. doi: 10.3233/JAD-230840

44. Zhou J, Liu W, Zhou H, Lau KK, Wong GHY, Chan WC, et al. Identifying dementia from cognitive footprints in hospital records among Chinese older adults: a machine-learning study. *Lancet Reg Health West Pac.* (2024) 46:101060. doi: 10.1016/j.lanwpc.2024.10106

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 12 June 2024
DOI 10.3389/fneur.2024.1379031

Check for updates

OPEN ACCESS

EDITED BY
Alejandro Rabinstein,
Mayo Clinic, United States

REVIEWED BY
Chang Liu,
Army Medical University, China
Nicolae Ovidiu Pop,
University of Oradea, Romania

*CORRESPONDENCE
Naibing Gu
✉ gunaibing@126.com

RECEIVED 30 January 2024
ACCEPTED 29 April 2024
PUBLISHED 12 June 2024

CITATION
Guo K, Zhu B, Li R, Xi J, Wang Q, Chen K, Shao Y, Liu J, Cao W, Liu Z, Di Z and Gu N (2024) Machine learning-based nomogram: integrating MRI radiomics and clinical indicators for prognostic assessment in acute ischemic stroke.
Front. Neurol. 15:1379031.
doi: 10.3389/fneur.2024.1379031

COPYRIGHT
© 2024 Guo, Zhu, Li, Xi, Wang, Chen, Shao, Liu, Cao, Liu, Di and Gu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Machine learning-based nomogram: integrating MRI radiomics and clinical indicators for prognostic assessment in acute ischemic stroke

Kun Guo[1], Bo Zhu[1], Rong Li[1], Jing Xi[1], Qi Wang[2], KongBo Chen[3], Yuan Shao[3], Jiaqi Liu[3], Weili Cao[1], Zhiqin Liu[1], Zhengli Di[1] and Naibing Gu[1]*

[1]Xi'an Central Hospital, Xi'an, China, [2]China-Japan Union Hospital of Jilin University, Changchun, China, [3]Tongchuan Mining Bureau Central Hospital, Tongchuan, China

Background: Acute Ischemic Stroke (AIS) remains a leading cause of mortality and disability worldwide. Rapid and precise prognostication of AIS is crucial for optimizing treatment strategies and improving patient outcomes. This study explores the integration of machine learning-derived radiomics signatures from multi-parametric MRI with clinical factors to forecast AIS prognosis.

Objective: To develop and validate a nomogram that combines a multi-MRI radiomics signature with clinical factors for predicting the prognosis of AIS.

Methods: This retrospective study involved 506 AIS patients from two centers, divided into training (n = 277) and validation (*n* = 229) cohorts. 4,682 radiomic features were extracted from T1-weighted, T2-weighted, and diffusion-weighted imaging. Logistic regression analysis identified significant clinical risk factors, which, alongside radiomics features, were used to construct a predictive clinical-radiomics nomogram. The model's predictive accuracy was evaluated using calibration and ROC curves, focusing on distinguishing between favorable (mRS $\leq$ 2) and unfavorable (mRS > 2) outcomes.

Results: Key findings highlight coronary heart disease, platelet-to-lymphocyte ratio, uric acid, glucose levels, homocysteine, and radiomics features as independent predictors of AIS outcomes. The clinical-radiomics model achieved a ROC-AUC of 0.940 (95% CI: 0.912–0.969) in the training set and 0.854 (95% CI: 0.781–0.926) in the validation set, underscoring its predictive reliability and clinical utility.

Conclusion: The study underscores the efficacy of the clinical-radiomics model in forecasting AIS prognosis, showcasing the pivotal role of artificial intelligence in fostering personalized treatment plans and enhancing patient care. This innovative approach promises to revolutionize AIS management, offering a significant leap toward more individualized and effective healthcare solutions.

KEYWORDS

acute ischemic stroke, radiomics, multi-parametric MRI, prognosis, nomogram, machine learning

Highlights

- High predictive accuracy for AIS prognosis.
- Integrates MRI radiomics with clinical factors.
- Utilizes advanced machine learning techniques.
- Provides a validated clinical-radiomics nomogram.
- Facilitates personalized AIS management.

1 Introduction

Ischemic stroke remains a formidable public health concern due to its high incidence, mortality, and morbidity rates, exerting a profound impact on society, families, and the affected individuals (1). Despite concerted efforts in recent years toward the management, treatment, and prevention of ischemic stroke, a significant proportion of patients fail to receive timely and effective intervention. This failure is often attributed to delayed recognition of symptoms, a lack of awareness regarding the urgency of medical care, and the unavailability of adequate facilities in primary healthcare settings, leading to varying extents of neurological deficits. The cornerstone of acute ischemic stroke treatment in the acute phase includes intravenous alteplase thrombolysis (2, 3) and mechanical thrombectomy (4). However, the application of intravenous thrombolysis is constrained by a narrow therapeutic time window, and stringent inclusion and exclusion criteria limit mechanical thrombectomy. Recent observations suggest a shift toward an increasing incidence of stroke among younger populations, a trend linked to improved living standards and heightened work-related stress (5, 6). Factors such as the timing of intervention, location and volume of the infarct, and post-stroke treatment and rehabilitation efforts are pivotal in determining patient outcomes and survival rates (7). Consequently, the accurate prediction of acute ischemic stroke prognosis becomes essential for evaluating the severity, identifying potential adverse outcomes, gauging rehabilitation prospects, and enhancing doctor-patient communication and clinical decision-making processes.

When cerebrovascular diseases are suspected or need exclusion, Computed Tomography Angiograms (CTA) and Magnetic Resonance Angiograms (MRA) have demonstrated high specificity and sensitivity. However, Digital Subtraction Angiograms (DSA) remains the definitive gold standard, providing unparalleled diagnostic insight. Despite its utility in detailing intravascular conditions, DSA's invasiveness and radiation exposure constrain its widespread clinical application.

Developing non-invasive methodologies with minimal radiation exposure is imperative in clinical practice to mitigate these limitations. Such advancements aim to improve the evaluation of treatment outcomes and prognostic accuracy in acute cerebral infarction.

Predictive models that amalgamate clinical observations, imaging findings, laboratory data, and other variables are instrumental in predictive assessment. These models enable comprehensive evaluations of rehabilitation prospects, survival rates, and disease incidence through mathematical and statistical approaches. A pioneering effort in this domain was conducted by Karen C. Johnston's team in 2000 (8), utilizing the NIH Stroke Scale (NIHSS), Barthel Index (BI), and Glasgow Coma Scale (GCS) to gauge acute ischemic stroke prognosis with promising results. Furthering this initiative, they integrated NIHSS scores and CT infarct volumes to adeptly predict patient outcomes at 3 months. Zhao et al. (9) further explored 30-day survival prediction in acute ischemic stroke patients by analyzing post-stroke blood routine and biochemical markers, including Neutrophil-to-Lymphocyte Ratio (NLR), Prognostic Nutritional Index (PNI), Systemic Immune-Inflammation Index (SII), and Risk Assessment (RA).

Radiomics, employing sophisticated image processing to extract detailed features from imaging studies of acute ischemic stroke patients, unveils in-depth insights into pathophysiological alterations. This technique enhances early diagnosis, disease type determination, precise lesion localization and quantification, and fosters accurate prognosis evaluation and treatment outcome assessment (10–13).

Clinical radiomics models, leveraging features derived from diffusion-weighted imaging (DWI), fluid-attenuated inversion recovery (FLAIR), and apparent diffusion coefficient (ADC) scans, have shown commendable efficacy in prognosticating outcomes for patients with acute ischemic stroke (14, 15). Despite these advancements, the utilization of imaging attributes and clinical data in appraising treatment effectiveness and forecasting the prognosis of acute ischemic stroke remains underexploited. Notably, a blend of T1-weighted images (T1w), T2-weighted images (T2w), and DWI is prevalently employed for assessing patients post-onset. This fact highlights the critical need for an exhaustive amalgamation of various imaging techniques to refine the precision and utility of predictive models in determining acute ischemic stroke outcomes. The potential to enhance predictive accuracy and clinical decision-making through such integrated models is vast yet underleveraged.

The objective of this research is to develop a model that effectively combines T1-weighted (T1w), T2-weighted (T2w), and diffusion-weighted imaging (DWI) features with pertinent clinical parameters. This model explores the associations between imaging characteristics and crucial clinical information, enhancing our understanding of acute ischemic stroke. The primary goal is to create an accurate and individualized decision support system that enriches the treatment process for patients experiencing acute ischemic stroke. We anticipate facilitating significantly improved patient outcomes by achieving this integration, ultimately benefiting those impacted by this condition.

2 Materials and methods

2.1 Subjects

This retrospective study recruited participants from two healthcare institutions, Xi'an Central Hospital (Center 1) and Tongchuan Mining Bureau Central Hospital (Center 2), with ethical approval from the respective hospitals' Ethics Committees by the Declaration of Helsinki. A collective cohort of 506 patients diagnosed with Acute Ischemic Stroke (AIS) was retrospectively analyzed across

both centers during the period from January to December 2021. Eligibility criteria for inclusion comprised admission within 24 h following symptom onset, an initial assessment using the National Institutes of Health Stroke Scale (NIHSS) upon admission, and undergoing diffusion-weighted imaging (DWI) within the first 72 h post-symptom onset.

Exclusion criteria were defined to omit patients who underwent reperfusion therapies for AIS, including intravenous thrombolysis with recombinant tissue plasminogen activator (rt-PA), urokinase (UK), and tenecteplase (TNK-tPA), as well as those who received bridging therapy (mechanical thrombectomy) or endovascular treatments (Center 1: $n = 17$, Center 2: $n = 31$). Given the unavailability of UK and TNK-tPA at the study sites, these treatments were not considered. Further exclusions applied to patients with hemorrhagic stroke, traumatic brain injury, subarachnoid hemorrhage, and hemorrhagic infarction (Center 1: $n = 1$, Center 2: $n = 2$), those presenting severe MRI artifacts (Center 1: $n = 1$, Center 2: $n = 5$), diagnosed with malignancies (Center 1: $n = 0$, Center 2: $n = 1$), or lost to follow-up (Center 1: $n = 10$, Center 2: $n = 41$; Figure 1).

Baseline clinical data, encompassing demographics, medical history (e.g., hypertension, hyperlipidemia, hyperuricemia, hypoproteinemia, hyperhomocysteinemia, diabetes, smoking, drinking, prior stroke, atrial fibrillation, coronary artery disease, chronic heart failure, arthrosis, hemadostenosis), TOAST classification, along with an extensive set of laboratory parameters and imaging features, were meticulously extracted from medical records. Two experienced neurologists, each with a decade of practice and blind to clinical and imaging data, conducted structured telephone interviews to determine patients' modified Rankin Scale (mRS) scores 6 months after hospital discharge, categorizing prognosis into favorable (mRS ≤ 2) and adverse (mRS > 2) outcomes. Although pre-stroke mRS scores are insightful for assessing baseline functionality, their exclusion aims to assure data integrity and minimize bias. The study's emphasis on uniformly assessed, quantifiable factors across all participants enhances the predictive model's validity, striving to eliminate confounding influences and bolster the reliability of the findings.

2.2 Image data acquisition

MRI scans were performed on all participants within 72 h post-symptom onset using the EXCITE HD 1.5 T MRI system by GE Healthcare, Milwaukee, WI, United States. Transverse T1-weighted fast spin echo (FSE) imaging was executed with specific parameters: TR/TE = 2,259/25.4 ms, slice thickness/gap = 5/1.5 mm, bandwidth = 244 Hz/Px, FOV = 240 × 240 mm^2, and acceleration factor (R) = 2. For T2-weighted FSE imaging, the settings were TR/TE = 5,582/111 ms, slice thickness/gap = 5/1.5 mm, bandwidth = 244 Hz/Px, FOV = 240 × 240 mm^2, and R = 3.

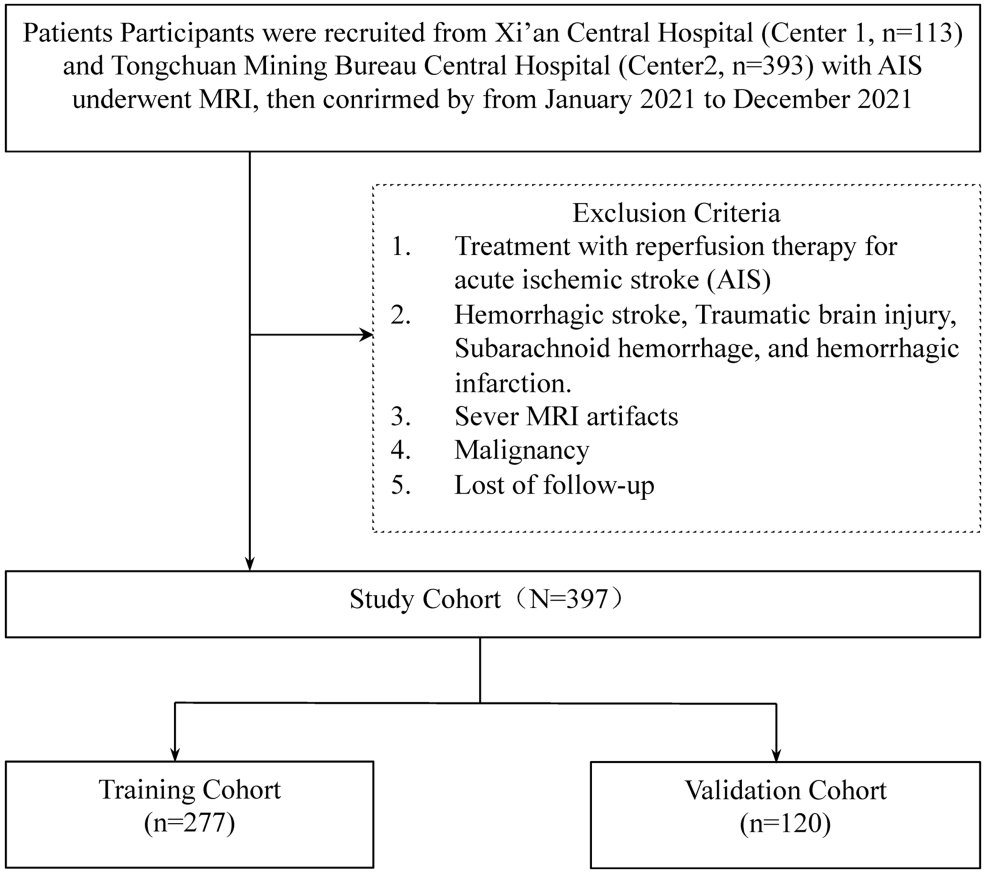


FIGURE 1
The workflow of the patient selection.

TABLE 1 Comparison of MRI sequence parameters.

MRI sequence	T1w	T2w	DWI
TR/TE(ms)	2,259/25.4	5,582/111	3,203/83.9
Slice Thickness/Gap(mm)	5/1.5	5/1.5	5/1.5
Bandwidth(Hz/Px)	244	244	3,906
FOV(mm)	240*240	240*240	240*240
Acceleration Factor(R)	2	3	2
b-values(s/mm)	/	/	0, 1,000

Diffusion-weighted imaging (DWI) utilized single-shot echo planar imaging (SS-EPI) with TR/TE = 3,203/83.9 ms, slice thickness/gap = 5/1.5 mm, bandwidth = 3,906 Hz/Px, FOV = 240 × 240 mm^2, R = 2, and b-values of 0 and 1,000 s/mm^2 (Table 1).

2.3 VOI delineated and radiomics feature extraction

Volumes of Interest (VOIs) were meticulously delineated using 3D-Slicer Software, while Pyradiomics software (version 3.0.1) facilitated the computation of radiomics features, adhering to the Image Biomarker Standardization Initiative's guidelines. A radiologist, blind to the patient's clinical information, performed the initial segmentation of MRI images. These segmentations were then reviewed and refined by a senior neuroradiologist with extensive experience. The segmentation process targeted the entire infarct region. Utilizing PyRadiomics, various radiomics features were extracted from these VOIs, including shape-based, first-order statistical, and several gray-level matrix features, from T1-weighted, T2-weighted, and diffusion-weighted images, totaling 4,682 features. This methodical extraction process ensures a comprehensive analysis of imaging data, which is crucial for evaluating acute ischemic stroke prognosis (Figure 2).

2.4 Features selection

Given the high-dimensional nature of radiomics features in acute ischemic stroke (AIS) analysis, the study aimed to pinpoint features significantly correlated with outcome predictions in the training cohort. Initial feature selection was performed using a U-test, setting a *p*-value threshold 0.05 to filter out non-significant and redundant features. Further refinement involved a correlation analysis to eliminate features with a correlation coefficient above 0.9. The Least Absolute Shrinkage and Selection Operator (LASSO) algorithm was then utilized to finalize feature selection, identifying the most predictive features through fivefold cross-validation.

2.5 Prediction development and diagnostic validation

Univariable logistic regression identified potential clinical predictors of AIS outcomes within the training cohort. Subsequently, significant predictors were analyzed through multivariable logistic regression, employing backward stepdown selection to isolate independent clinical predictors. These findings were presented as odds ratios with corresponding 95% confidence intervals, forming the basis of a multivariable clinical prediction model. The model combined these independent clinical predictors with the radiomics signature to create a comprehensive clinical-radiomics model, which underwent rigorous evaluation through ROC curve analysis and other statistical measures to assess its discriminative performance.

2.6 Clinical usefulness and calibration curves of the clinical-radiomics model

The clinical-radiomics model's calibration was examined using calibration curves alongside the Hosmer-Lemeshow test to determine the model's fit accuracy.

2.7 Statistics analysis

The study used statistical methods to analyze demographic and clinical data, including independent *t*-tests for normally distributed data, Mann–Whitney U tests for non-normally distributed data, and chi-square tests for categorical variables. The predictive performance of clinical, radiomics, and clinical-radiomics models was evaluated through Receiver Operating Characteristic (ROC) curves. Model comparisons were conducted using the Delong test. Calibration of the clinical-radionics model was assessed with calibration curves and the Hosmer-Lemeshow test. Statistical analyses were conducted in R software, with significance determined by a two-tailed *p*-value of <0.05.

3 Results

3.1 Baseline characteristics of patients

The baseline characteristics of patients within both the training and validation groups are detailed in Table 2. Analysis revealed no statistically significant disparities between the cohorts ($p > 0.05$), indicating comparable baseline conditions. Within the training cohort, patients experiencing unfavorable outcomes (modified Rankin Scale, mRS, > 2) accounted for 66.78% (182 out of 277), while in the validation group, this proportion stood at 60.83% (73 out of 120).

3.2 Radiomics feature selection and LASSO logistic regression findings

In refining our dataset, redundant features were eliminated through a U-test and Spearman correlation analysis, narrowing down the selection from an initial pool of 4,682 image features extracted from the Volumes of Interest (VOIs) to 791 radiomic features. This step was essential for enhancing the dataset's manageability and relevance. Following this initial reduction, the Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression method was applied to further distill these features, focusing on identifying those of optimal predictive value. Determining the most appropriate lambda value was crucial in this context; to this end, five-fold

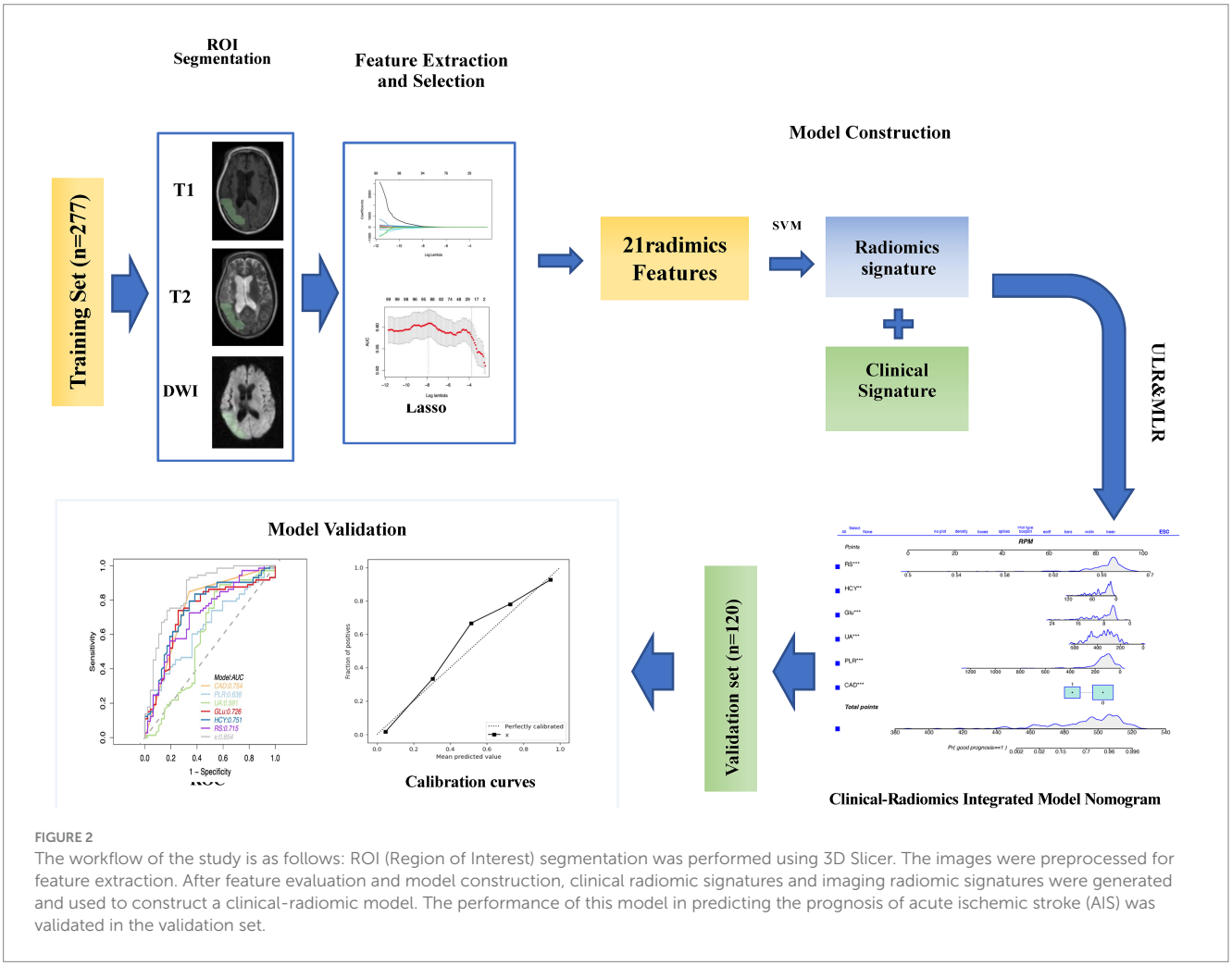


FIGURE 2
The workflow of the study is as follows: ROI (Region of Interest) segmentation was performed using 3D Slicer. The images were preprocessed for feature extraction. After feature evaluation and model construction, clinical radiomic signatures and imaging radiomic signatures were generated and used to construct a clinical-radiomic model. The performance of this model in predicting the prognosis of acute ischemic stroke (AIS) was validated in the validation set.

TABLE 2 The characteristic of 21 final feature for radiomics assessment.

Radiomic features
T1_original_shape_SurfaceArea
T1_original_firstorder_Energy
T1_wavelet.LLH_ngtdm_Coarseness
T1_wavelet.LHL_glcm_Idmn
T1_wavelet.LHH_glcm_Idmn
T1_wavelet.HLH_glcm_ClusterShade
T1_wavelet.HHL_glrlm_LongRunEmphasis
T1_exponential_glszm_ZoneEntropy
T1_exponential_gldm_DependenceEntropy
T1_squareroot_glcm_ClusterShade
T1_gradient_glrlm_LongRunEmphasis
T2_original_ngtdm_Strength
T2_wavelet.LLL_firstorder_90Percentile
T2_wavelet.LLL_firstorder_InterquartileRange
T2_square_ngtdm_Strength
DWI_log.sigma.1.0.mm.3D_glcm_DifferenceVariance
DWI_wavelet.LHL_firstorder_Mean
DWI_wavelet.LHL_firstorder_Median
DWI_wavelet.LHL_glcm_InverseVariance
DWI_wavelet.HLH_firstorder_Skewness
DWI_wavelet.HHH_firstorder_Kurtosis

cross-validation was utilized, selecting a lambda value within one standard error of the minimum. Through meticulous selection, a final set of 21 radiomic features was identified (as detailed in Table 3). This rigorous methodology underscores the accuracy and efficacy of the resulting predictive model (illustrated in Figure 3).

3.3 Establishment and performance of the clinical features

Table 4 presents the results of the multivariate logistic regression analysis, identifying significant variables ($p < 0.05$) such as Coronary Artery Disease (CAD), White Blood Cell count (WBC), Platelet-to-Lymphocyte Ratio (PLR), Uric Acid (UA), Glucose (Glu), Homocysteine (HCY), and the Radiomics Score (RS). These variables have been pinpointed as independent predictors for clinical functional outcomes. Utilizing these determinants, we constructed a radiomics nomogram within the training set, creating a clinical-radiomics comprehensive prediction model (illustrated in Figure 4).

The nomogram is meticulously designed, highlighting the relative importance of variables identified through multivariate regression analysis. A scale is displayed at the top of the column chart, with predictive model variables listed on the left. Each independent variable is assigned a score, correlating to specific values, with the length of each segment indicating its contributory weight to the outcome event. Hence, longer segments underscore a

TABLE 3 Baseline characteristics of patients in the training and validation cohorts.

	Training cohort (*n* = 277)			Validation cohort (*n* = 120)		
	Poor (mRS > 2) *n* = 95	Good (mRS ≤ 2) *n* = 182	p.overall	Poor (mRS > 2) *n* = 47	Good (mRS ≤ 2) *n* = 73	p.overall
Age	68.0 [58.5;77.5]	68.0 [60.0;75.0]	0.477	71.0 [62.5;81.0]	69.0 [64.0;78.0]	0.647
Gender:			0.824			0.146
Female	18 (18.9%)	38 (20.9%)		19 (40.4%)	19 (26.0%)	
Male	77 (81.1%)	144 (79.1%)		28 (59.6%)	54 (74.0%)	
Hypertension:			0.003			0.013
No	10 (10.5%)	49 (26.9%)		3 (6.38%)	19 (26.0%)	
Yes	85 (89.5%)	133 (73.1%)		44 (93.6%)	54 (74.0%)	
Hyperlipemia:			<0.001			0.001
No	14 (14.7%)	121 (66.5%)		15 (31.9%)	48 (65.8%)	
Yes	81 (85.3%)	61 (33.5%)		32 (68.1%)	25 (34.2%)	
Hyperuricemia:			<0.001			<0.001
No	22 (23.2%)	150 (82.4%)		22 (46.8%)	62 (84.9%)	
Yes	73 (76.8%)	32 (17.6%)		25 (53.2%)	11 (15.1%)	
Hypoproteinemia:			0.052			0.066
No	77 (81.1%)	164 (90.1%)		35 (74.5%)	65 (89.0%)	
Yes	18 (18.9%)	18 (9.89%)		12 (25.5%)	8 (11.0%)	
Hyperhomocysteinemia:			<0.001			0.04
No	7 (7.37%)	65 (35.7%)		10 (21.3%)	30 (41.1%)	
Yes	88 (92.6%)	117 (64.3%)		37 (78.7%)	43 (58.9%)	
Diabetes:			<0.001			<0.001
No	21 (22.1%)	127 (69.8%)		14 (29.8%)	57 (78.1%)	
Yes	74 (77.9%)	55 (30.2%)		33 (70.2%)	16 (21.9%)	
Smoking history:			<0.001			<0.001
No	29 (30.5%)	140 (76.9%)		26 (55.3%)	65 (89.0%)	
Yes	66 (69.5%)	42 (23.1%)		21 (44.7%)	8 (11.0%)	
Drinking history:			<0.001			<0.001
No	42 (44.2%)	166 (91.2%)		27 (57.4%)	71 (97.3%)	
Yes	53 (55.8%)	16 (8.79%)		20 (42.6%)	2 (2.74%)	
Stroke history:			0.12			0.135
No	49 (51.6%)	113 (62.1%)		19 (40.4%)	41 (56.2%)	
Yes	46 (48.4%)	69 (37.9%)		28 (59.6%)	32 (43.8%)	
AF*			1			1
No	87 (91.6%)	168 (92.3%)		44 (93.6%)	69 (94.5%)	
Yes	8 (8.42%)	14 (7.69%)		3 (6.38%)	4 (5.48%)	
CAD*			<0.001			<0.001
No	30 (31.6%)	146 (80.2%)		16 (34.0%)	62 (84.9%)	
Yes	65 (68.4%)	36 (19.8%)		31 (66.0%)	11 (15.1%)	
AS*			<0.001			0.013
No	18 (18.9%)	95 (52.2%)		10 (21.3%)	33 (45.2%)	
Yes	77 (81.1%)	87 (47.8%)		37 (78.7%)	40 (54.8%)	

(Continued)

TABLE 3 (Continued)

	Training cohort (*n* = 277)			Validation cohort (*n* = 120)		
	Poor (mRS > 2) *n* = 95	Good (mRS ≤ 2) *n* = 182	p.overall	Poor (mRS > 2) *n* = 47	Good (mRS ≤ 2) *n* = 73	p.overall
Hemadostenosis:			<0.001			0.006
No	20 (21.1%)	105 (57.7%)		10 (21.3%)	35 (47.9%)	
Yes	75 (78.9%)	77 (42.3%)		37 (78.7%)	38 (52.1%)	
ADL*	55.0 [40.0;60.0]	80.0 [65.0;100]	<0.001	55.0 [42.5;62.5]	80.0 [65.0;100]	<0.001
NHISS*	7.00 [5.00;9.50]	2.00 [1.00;3.00]	<0.001	7.00 [5.00;11.5]	2.00 [1.00;3.00]	<0.001
TOAST*			0.007			0.636
Large artery atherosclerosis	8 (8.42%)	7 (3.85%)		0 (0.00%)	2 (2.74%)	
Cardioembolism	72 (75.8%)	138 (75.8%)		39 (83.0%)	60 (82.2%)	
Small vessel occlusion	9 (9.47%)	35 (19.2%)		4 (8.51%)	8 (11.0%)	
Other determined etiology	6 (6.32%)	2 (1.10%)		3 (6.38%)	3 (4.11%)	
Undetermined etiology				1 (2.13%)	0 (0.00%)	
KWST*			<0.001			0.001
0	1 (1.05%)	4 (2.20%)		0 (0.00%)	2 (2.74%)	
1	59 (62.1%)	162 (89.0%)		33 (70.2%)	67 (91.8%)	
2	24 (25.3%)	11 (6.04%)		6 (12.8%)	3 (4.11%)	
3	7 (7.37%)	3 (1.65%)		5 (10.6%)	0 (0.00%)	
4	2 (2.11%)	1 (0.55%)		3 (6.38%)	1 (1.37%)	
5	2 (2.11%)	1 (0.55%)				
WBC	10.9 [8.39;12.7]	6.78 [5.56;8.34]	<0.001	11.2 [8.09;13.4]	6.63 [5.55;8.49]	<0.001
NEU	5.53 [4.18;7.18]	4.94 [3.62;6.43]	0.116	6.71 [4.72;8.57]	4.99 [3.61;6.62]	0.005
LYM	0.99 [0.89;1.69]	1.71 [1.30;2.19]	<0.001	1.03 [0.87;1.50]	1.57 [1.24;2.18]	<0.001
NLR*	4.73 [2.85;6.42]	2.83 [1.90;4.05]	<0.001	6.03 [3.40;9.48]	3.22 [1.97;4.73]	<0.001
MON	0.43 [0.33;0.64]	0.47 [0.35;0.62]	0.104	0.45 [0.35;0.58]	0.44 [0.35;0.56]	0.796
MLR*	0.37 [0.24;0.50]	0.28 [0.20;0.36]	<0.001	0.45 [0.30;0.65]	0.28 [0.20;0.39]	0.001
Hb	145 [135;153]	143 [132;154]	0.468	140 [132;149]	145 [133;156]	0.285
HCT	0.44 [0.41;0.50]	0.44 [0.41;3.69]	0.629	0.42 [0.40;0.45]	0.43 [0.40;0.49]	0.541
PLT	179 [148;219]	178 [147;220]	0.772	180 [130;204]	189 [133;219]	0.292
SII*	842 [441;1,287]	459 [298;686]	<0.001	858 [534;1,750]	548 [280;895]	0.002
PLR*	150 [98.8;212]	103 [71.9;143]	<0.001	138 [102;228]	115 [74.2;168]	0.011
TBIL	17.8 [13.4;24.0]	16.4 [12.9;21.5]	0.257	19.1 [15.9;26.4]	15.7 [12.3;21.4]	0.011
Alb	38.7 [36.5;42.2]	39.3 [37.3;42.5]	0.309	38.9 [36.7;40.9]	38.9 [36.2;41.8]	0.828
Glb	22.0 [20.0;26.8]	22.4 [19.4;27.3]	0.9	22.5 [19.9;26.0]	22.8 [20.2;26.2]	0.94
ALT	17.0 [12.0;23.0]	16.5 [12.0;22.8]	0.515	16.0 [12.0;23.5]	17.0 [10.0;21.0]	0.735
AST	20.0 [17.0;24.5]	18.5 [16.0;23.0]	0.245	19.0 [15.0;26.5]	18.0 [16.0;23.0]	0.577
AST/ALT	1.38 [1.00;2.25]	1.73 [1.10;2.31]	0.072	1.88 [1.09;2.48]	2.00 [1.20;2.62]	0.55
Urea	5.21 [4.46;6.23]	5.26 [4.48;6.22]	0.873	5.68 [4.73;6.61]	5.78 [4.76;6.75]	0.919
Cr	67.0 [56.5;77.0]	66.6 [55.0;77.8]	0.783	64.0 [57.0;74.6]	63.0 [57.0;74.0]	0.906
UA	409 (104)	324 (88.2)	<0.001	362 (115)	327 (85.9)	0.074
TC	4.35 [3.51;5.38]	4.06 [3.50;4.94]	0.15	4.35 (1.11)	4.40 (1.11)	0.823
PNI*	44.9 [42.2;49.3]	47.9 [43.4;52.3]	0.016	45.0 [41.9;49.0]	49.0 [41.8;54.2]	0.09

(Continued)

TABLE 3 (Continued)

	Training cohort (*n* = 277)			Validation cohort (*n* = 120)		
	Poor (mRS > 2) *n* = 95	Good (mRS ≤ 2) *n* = 182	p.overall	Poor (mRS > 2) *n* = 47	Good (mRS ≤ 2) *n* = 73	p.overall
TG	1.97 [1.05;3.12]	1.35 [0.99;1.95]	0.001	1.34 [0.92;2.42]	1.25 [0.93;1.74]	0.154
HDL	0.98 [0.82;1.15]	1.05 [0.91;1.28]	0.005	1.12 [0.92;1.27]	1.09 [0.96;1.31]	0.807
LDL	2.89 [2.18;3.48]	2.31 [1.70;2.92]	<0.001	2.36 [1.79;3.42]	2.44 [1.92;3.18]	0.517
Glu	11.2 [7.25;13.9]	5.68 [5.04;7.74]	<0.001	11.9 [6.66;13.6]	5.86 [5.14;6.98]	<0.001
HCY	45.2 [24.7;56.2]	19.6 [14.2;29.4]	<0.001	43.7 [19.2;54.5]	16.7 [13.1;22.9]	<0.001
K	3.96 (0.38)	4.08 (0.44)	0.018	4.05 (0.40)	4.09 (0.42)	0.564
Na	141 [139;143]	141 [139;143]	0.532	142 [140;143]	141 [139;143]	0.833
Cl	106 [103;108]	106 [104;108]	0.402	106 [104;108]	105 [103;108]	0.739
Ca	2.25 (0.14)	2.25 (0.14)	0.903	2.25 [2.13;2.33]	2.23 [2.15;2.32]	0.899
P	0.98 (0.19)	0.99 (0.21)	0.595	0.98 (0.22)	0.99 (0.22)	0.859
PT	10.8 [10.2;11.4]	10.9 [10.2;11.6]	0.683	11.0 [10.3;11.9]	10.7 [10.2;11.4]	0.359
INR	0.95 [0.89;1.00]	0.95 [0.89;1.03]	0.663	0.97 [0.90;1.08]	0.95 [0.88;0.99]	0.121
APTT	26.0 [24.4;28.9]	26.8 [24.0;30.4]	0.438	25.9 [23.1;30.2]	26.2 [23.9;28.9]	0.821
TT	16.9 [15.2;17.9]	16.6 [15.2;17.9]	0.489	16.6 [15.4;17.9]	16.8 [15.9;17.6]	0.604
FIB	2.85 [2.33;3.38]	2.82 [2.35;3.36]	0.832	2.61 [2.27;3.36]	2.66 [2.22;3.20]	0.526
D-Dimer	0.46 [0.23;1.17]	0.34 [0.21;0.64]	0.04	0.56 [0.29;1.17]	0.40 [0.26;0.69]	0.089
FDP	1.60 [1.00;3.60]	1.36 [0.90;2.30]	0.06	1.50 [1.13;3.20]	1.30 [0.70;2.20]	0.08
EF	57.0 [55.0;65.0]	58.0 [55.0;64.0]	0.996	61.0 [55.5;65.0]	60.0 [55.0;65.0]	0.363
LEVF	110 [101;124]	115 [105;131]	0.078	114 [104;125]	113 [105;127]	0.815
LVDD*	30.0 [28.0;32.0]	30.0 [28.0;32.0]	0.66	30.0 [28.0;32.0]	30.0 [28.0;33.0]	0.462
LVSD*	44.0 [41.0;46.0]	45.0 [43.0;48.0]	0.012	45.0 [42.0;47.0]	45.0 [43.0;48.0]	0.599
HR	76.0 [69.5;81.5]	76.0 [67.0;82.0]	0.704	77.0 [71.0;80.0]	75.0 [66.0;82.0]	0.344
SBP	150 [138;162]	145 [133;155]	0.03	151 (21.1)	150 (19.8)	0.879
DBP	88.0 [78.0;95.5]	86.0 [78.0;94.8]	0.3	85.0 [78.5;96.0]	86.0 [76.0;93.0]	0.743
Ht	170 [160;175]	170 [165;174]	0.717	170 [160;172]	168 [162;172]	0.901
IdealWt	64.2 [57.0;68.0]	65.0 [59.8;66.5]	0.97	62.0 [56.0;66.1]	63.5 [57.8;65.8]	0.696
Wt	65.0 [60.0;74.0]	66.5 [61.2;74.0]	0.894	65.0 [56.5;70.0]	65.0 [55.0;70.0]	0.859
BMI	23.8 [21.5;25.4]	23.2 [22.0;25.0]	0.251	22.6 [20.7;25.2]	23.0 [20.3;24.8]	0.866
GNRI*	62.2 [58.6;67.1]	63.1 [59.8;67.6]	0.288	62.4 [58.8;65.6]	62.4 [57.5;67.3]	0.782
GNRI grade:						
1	0 (0.00%)	1 (0.55%)				
3	95 (100%)	181 (99.5%)		47 (100%)	73 (100%)	
RS*	0.66 [0.65;0.67]	0.67 [0.66;0.67]	<0.001	0.66 [0.65;0.67]	0.67 [0.66;0.68]	<0.001

AF, atrial fibrillation; CAD, coronary artery disease; AS, atherosclerosis; ADL, activities of daily living; NIHSS, National Institutes of Health Stroke Scale; TOAST, Trial of Org 10172 in Acute Stroke Treatment; KWST, Kubota Water Swallowing Test; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; SII, systemic immune-inflammation index; PLR, platelet-to-lymphocyte ratio; PNI, prognostic nutritional index; LVDD, left ventricular end-diastolic diameter; LVSD, left ventricular end-systolic diameter; GNRI, geriatric nutritional risk index; RS, radiomic signature. Data for categorical variables are expressed in terms of the number of patients (percentage), and continuous variables are represented by median [interquartile range]. The *p*-value is used for group comparisons.

variable's heightened significance. By aggregating scores for individual variables, a total score is achievable. This cumulative score facilitates the determination of the linear predictor and predicted probability for Acute Ischemic Stroke (AIS) prognosis through the scale provided below. A prediction value of 1 suggests a favorable prognosis, whereas 0 implies a less favorable prognosis.

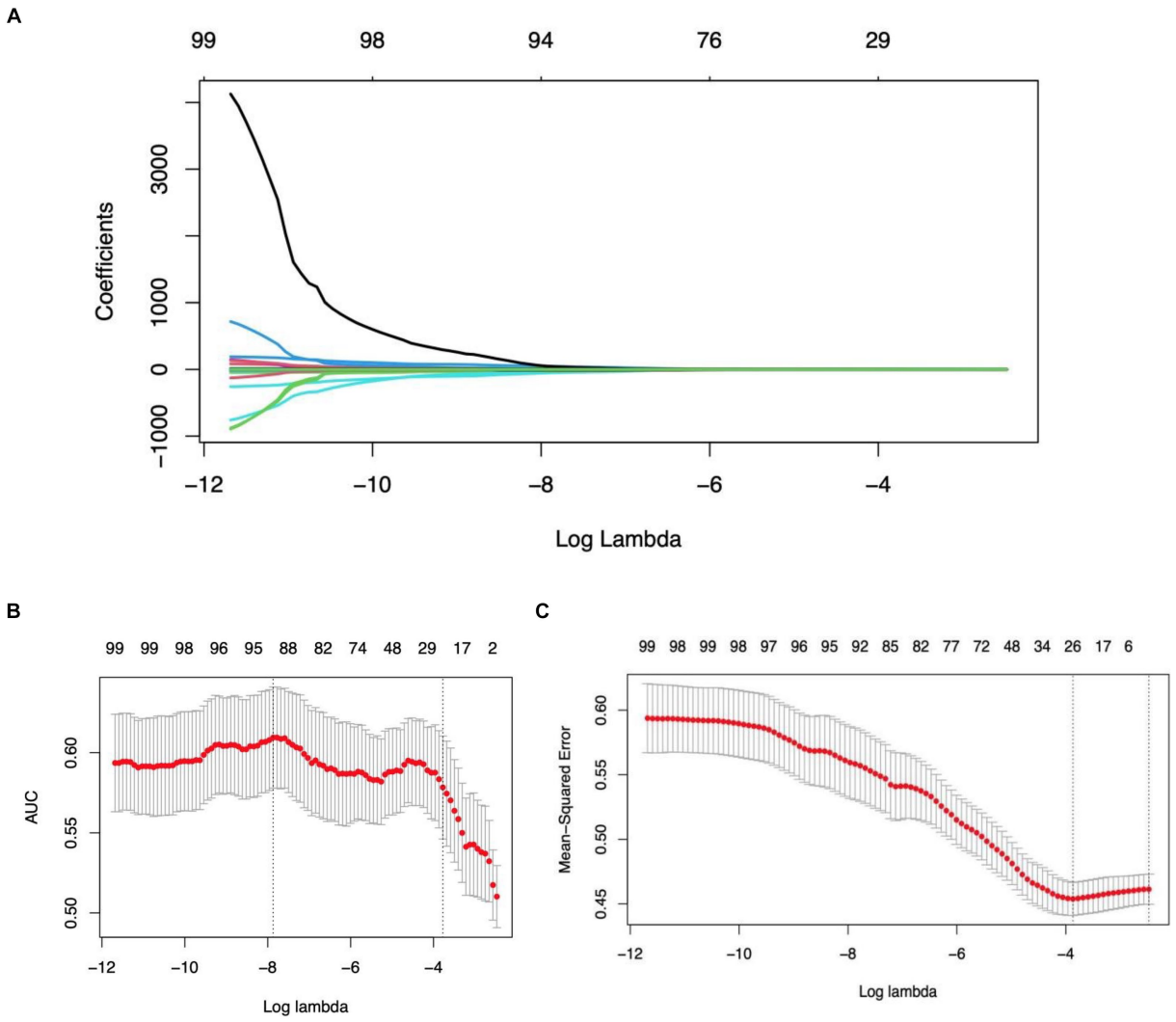


FIGURE 3
Illustration of the LASSO model's application, employing a tuning parameter (lambda) and utilizing five-fold cross-validation based on both the minimum criteria and one standard error (1se), facilitating the selection of radiomics features during the feature selection phase.

TABLE 4 Multivariate logistic regression.

Variable	OR(95%CI)	*p*-value	
CAD	0.108(0.043, 0.255)	<0.001	***
PLR	0.991 (0.985, 0.996)	<0.001	***
UA	0.992 (0.988, 0.995)	<0.001	***
Glu	0.789 (0.709, 0.874)	<0.001	***
HCY	0.971 (0.954, 0.989)	<0.001	***
RS	1.86e + 38 (9.48e + 24, 3.30e + 55)	<0.001	***

Significant codes: 0; "***"0.001; "**"0.01; "*"0.05;"."0.1; " "1.

3.4 Performance of the combined clinical-radiomics model

The performance of the combined clinical-radiomics model was meticulously evaluated using Receiver Operating Characteristic (ROC) curves. This comprehensive model exhibited impressive predictive accuracy for Acute Ischemic Stroke (AIS) outcomes, as evidenced by Area Under the Curve (AUC) scores of 0.940 (95% Confidence Interval [CI]: 0.912–0.969) in the training cohort and 0.853 (95% CI: 0.781–0.926) in the validation cohort. Notably, in the validation cohort, the clinical-radiomics model achieved high sensitivity (91.8%) and moderate specificity (68.1%), indicating its robust ability to accurately predict patient outcomes. The Positive Predictive Value (PPV) of the clinical-radiomics approach in the validation set was notably high, at approximately 81.7%. The Radiomics Score (RS) and the clinical-radiomics models yielded AUC values of 0.715 (95% CI: 0.619–0.810) and 0.854 (95% CI: 0.781–0.926), respectively, as illustrated in Figure 5.

Clinical variables such as Coronary Artery Disease (CAD), Platelet-to-Lymphocyte Ratio (PLR), Uric Acid (UA), Glucose (Glu), and Homocysteine (HCY) demonstrated significant predictive utility, as detailed in Table 5. The DeLong test revealed no significant differences in the performance of these three individual models (all $p > 0.05$), underscoring the consistency of the predictive capability across the models.

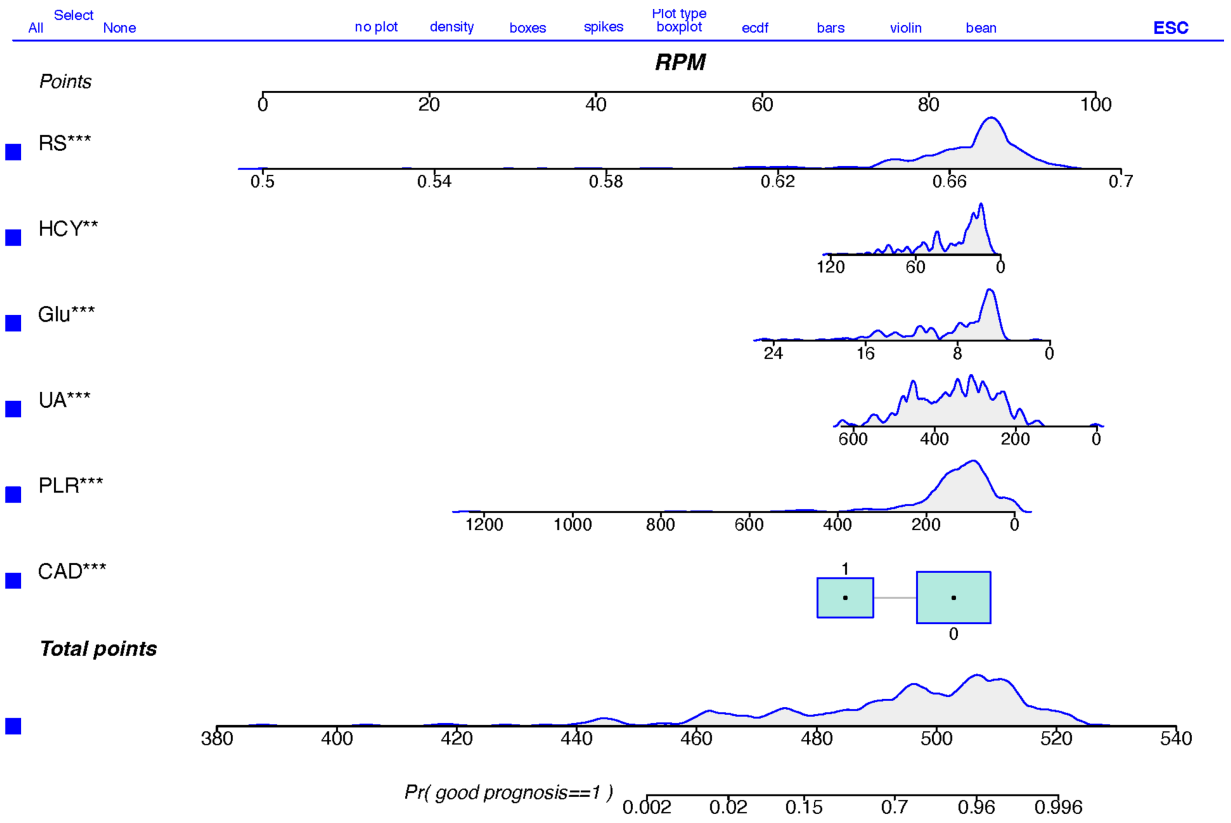


FIGURE 4
A nomogram based on clinical-radiomics models for predicting AIS outcomes.

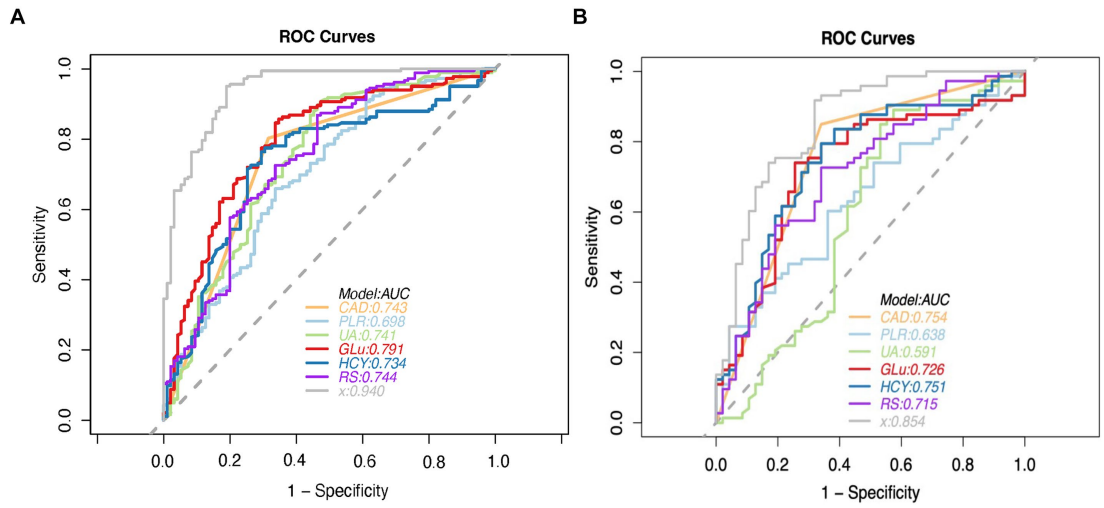


FIGURE 5
ROC curve analysis of predictive models for modified rankin scale (MRS) outcomes in the training **(A)** and validation **(B)** cohorts, with the clinical-radiomics model highlighted. CAD = coronary artery disease, PLR = platelet-lymphocyte ratio, UA = serum uric acid level, GLU = fasting plasma glucose, HCY = homocysteine, RS = Radiomics Signature, x = Clinical-Radiomics model.

3.5 Clinical usefulness and calibration curves for the clinical-radiomics

When contrasted with single-scale prediction models, the clinical-radiomics model showcased a superior capability in discriminating performance evaluation. As depicted in Figure 6, calibration plots demonstrated a commendable congruence between the model's predictions and the actual clinical outcomes for Acute Ischemic Stroke (AIS). This alignment indicates that the clinical-radiomics model provides a significantly enhanced net benefit over traditional single-scale models, underscoring its greater clinical utility.

TABLE 5 Predictive performance of three models in the training and validation cohorts.

Model	Training cohort (n = 277)					
	AUC (95% CI*)	Sensitivity	Specificity	Accuracy	PPV*	NPV*
CAD	0.743(0.688-0.798)	0.802	0.684	0.762	0.83	0.643
PLR	0.698(0.631–0.765)	0.659	0.663	0.66	0.789	0.504
UA	0.741(0.676–0.806)	0.907	0.526	0.776	0.786	0.746
Glu	0.791(0.734–0.849)	0.857	0.653	0.787	0.825	0.705
HCY	0.734(0.669–0.796)	0.774	0.694	0.747	0.829	0.617
Radiomics model	0.744(0.681–0.808)	0.868	0.536	0.755	0.782	0.68
Clinical-radiomics model	0.940(0.912–0.969)	0.951	0.811	0.903	0.906	0.895
Model	Validation cohort (n =120)					
	AUC (95% CI)	Sensitivity	Specificity	Accuracy	PPV	NPV
CAD	0.754(0.675–0.834)	0.85	0.66	0.775	0.795	0.738
PLR	0.591(0.538–0.738)	0.603	0.638	0.617	0.721	0.508
UA	0.726(0.477–0.705)	0.468	0.849	0.7	0.713	0.667
Glu	0.791(0.630–0.822)	0.74	0.745	0.742	0.818	0.648
HCY	0.751(0.659–0.842)	0.795	0.66	0.742	0.784	0.674
Radiomics model	0.714(0.619–0.810)	0.726	0.66	0.7	0.768	0.608
Clinical-radiomics model	0.854(0.781–0.926)	0.918	0.681	0.825	0.817	0.842

*CI, confidence interval; PPV, Positive predictive value; NPV, Negative predictive value.

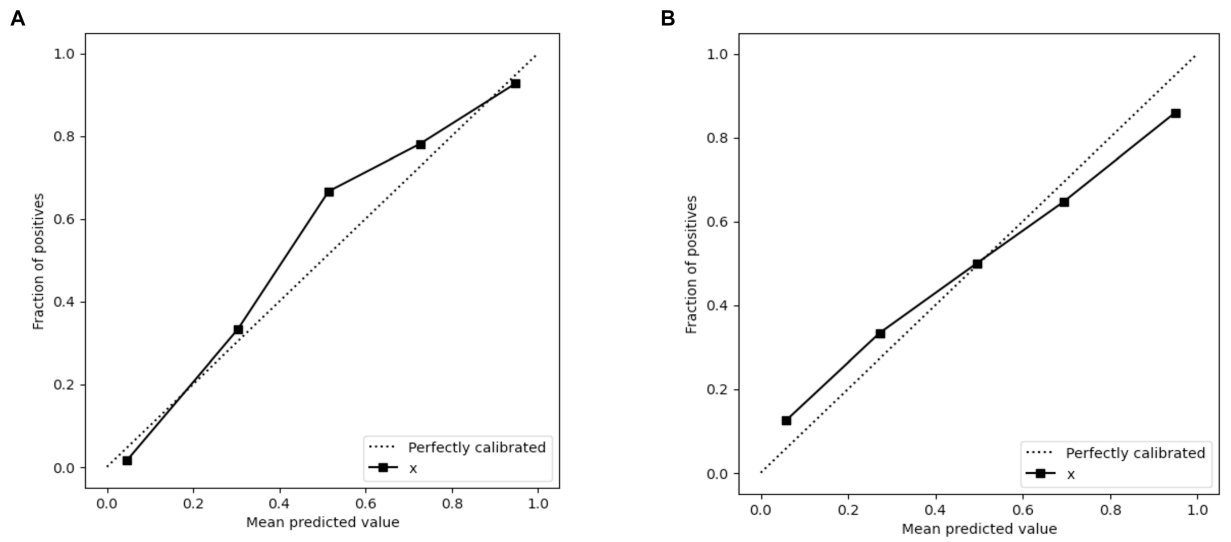


FIGURE 6
Calibration curves of the clinical-radiomics model in the training set **(A)** and validation set **(B)** for predicting AIS outcomes. x = Clinical-Radiomics model.

4 Discussion

Clinical predictive models are increasingly utilized in acute ischemic stroke management, serving as a critical role in prognosis and diagnostic support. Such as the widely recognized modified Rankin Scale (mRS), which could provide prognostic insights based on thorough patient evaluations, have gained broad acceptance in clinical settings. Historically, studies on ischemic stroke prognosis have predominantly utilized non-imaging data, employing various statistical models and machine learning algorithms such as linear regression, support vector machine (SVM), k-nearest neighbors (KNN), logistic regression, decision trees, k-means clustering, random forests (RF), naive Bayes, dimensionality reduction techniques, and gradient boosting. However, these methods have limitations,

particularly their dependence on basic clinical data without incorporating detailed biomarkers or complex imaging findings. Such limitations may impede a comprehensive understanding of the multifaceted nature of the disease. Moreover, despite the importance of these thorough assessments, they may not adequately account for individual variability and specific lesion characteristics inherent in the nonlinear dynamics and subjective assessments of stroke prognosis evaluation systems, posing challenges in reflecting the dynamic progression of the condition (16, 17). The advancement of precision medicine calls for more sophisticated and precise methods capable of navigating the intricacies of ischemic stroke prognosis. In this context, artificial intelligence (AI) models emerge as potent tools, offering the potential for more tailored and accurate prognostic predictions.

The progression of precision medicine necessitates more sophisticated and precise methodologies capable of addressing the intricate nature of ischemic stroke prognosis. As a result, there is a growing focus on integrating diverse data sources, such as clinical observations, radiomics features, and biomarkers, to construct predictive models that provide improved accuracy, comprehensiveness, and personalization. Technological advancements have facilitated an increase in studies integrating imaging data (e.g., CT and MRI scans) with state-of-the-art artificial intelligence machine learning techniques to enhance the accuracy and sensitivity of stroke prognosis prediction models. This evolving paradigm highlights the comprehensive utilization of diverse data types, especially imaging data, creating new opportunities for advancing stroke prognosis research and achieving more detailed and precise assessments of patient prognoses in the medical domain (18).

Radiomics and artificial intelligence hold great promise in the management of acute stroke. Cranial CT is the standard screening tool post-stroke, with the Alberta Stroke Program Early CT score (ASPECTS) being a rapid, straightforward, and reliable method for assessing early ischemic changes in patients with ischemic stroke, which is crucial for predicting treatment outcomes and prognosis (19). Nicolae et al. (20) retrospectively analyzed 340 patients revealed that ASPECTS effectively predicts the prognosis of patients with acute ischemic stroke (AIS), with lower scores indicating a larger infarct volume, particularly in diabetic and elderly patients. Chen et al. (21) retrospectively analyzed the CT images of 276 AIS patients, confirming the reliability and accuracy of automated ASPECTS scoring software. Hulin Kuang et al. (22) introduced EIS-Net, a novel multi-task learning network capable of simultaneously segmenting early infarcts and scoring ASPECTS on non-contrast CT images, offering performance comparable to expert assessments in a rapid manner. Masaki et al. (23) developed a deep learning-based automated ASPECTS calculation software utilizing the 3D-BHCA algorithm, demonstrating higher accuracy and efficiency than traditional methods, which could assist physicians in formulating superior treatment plans. Additionally, Qi et al. (24) explored apparent diffusion coefficient (ADC) image signal changes and their quantitative assessments across 207 acute ischemic stroke patients. Their findings suggested these analyses could act as crucial references for estimating acute ischemic stroke volume, thus offering valuable diagnostic insights for clinical practice. Ma's study (25) corroborated these observations. Future research may integrate multimodal imaging with clinical factors, offering more possibilities for precision diagnosis.

Cerebrovascular malformations, commonly arising from the abnormal development of intracranial vessels, can precipitate severe events such as stroke. Moyamoya disease (MMD), predominantly affecting children and adults, is marked by arterial stenosis and occlusion, with the potential to cause conditions like epilepsy and cognitive delays in pediatric patients. Despite advancements in modern imaging facilitating diagnosis, therapeutic options for MMD are limited, underscoring the critical importance of early detection and intervention to improve patient outcomes (26). The emergence of artificial intelligence (AI), particularly in the realms of deep learning and machine learning, has introduced novel approaches to MMD diagnosis.

In Kim's study (27), they utilized deep learning and convolutional neural network (CNN) techniques to analyze cranial images from 345 diagnosed MMD patients and 408 control subjects for the detection of MMD. CNN for the analysis of images from MMD patients and controls demonstrated impressive levels of accuracy, sensitivity, and specificity, underscoring the potential of AI in medical imaging.

Qin et al. (28) utilized machine learning models to analyze DSA images and predict mean transit time in MMD or Moyamoya syndrome (MMS) patients, achieving high accuracy in specific brain regions, which could offer significant insights for clinical diagnosis and treatment. Collectively, these studies suggest a promising role for AI in refining MMD diagnosis and treatment strategies.

Multimodal MRI, incorporating a variety of imaging sequences such as DWI, FLAIR, susceptibility-weighted imaging (SWI), and T1w and T2w, offers a rich source of biological insights. This comprehensive imaging approach enables physicians to thoroughly understand the characteristics and extent of lesions, thereby enhancing the precision of disease diagnoses. Additionally, employing radiomics and machine learning techniques, multimodal MRI facilitates the automated extraction and analysis of extensive imaging features. By amalgamating data from different MRI sequences, this strategy supports the development of predictive models capable of accurately determining patient outcomes. The confluence of multiple imaging modalities with sophisticated analytical methods promises more accurate and personalized medical evaluations.

In another significant effort, Quan et al. (14) derived radiomics features from fluid-attenuated inversion recovery (FLAIR) and ADC images across 753 cases, demonstrating their utility in forecasting clinical outcomes in acute ischemic stroke (AIS) patients. External validation showed promising area-under-the-curve (AUC) values across various models, with the combined ADC and FLAIR radiomics model significantly enhancing predictive accuracy for adverse outcomes.

A meta-analysis by Hanna Maria Dragoş et al. (29), reviewing 150 articles, suggests models that merge clinical and imaging features more effectively predict disability outcomes in stroke patients at 3 and 6 months post-event. Wang's study (30) established three radiomics models and a comprehensive nomogram that integrated clinical features and radiomic signatures, showing robust predictive performance for post-thrombolytic ischemic stroke prognosis.

To predict ischemic stroke outcomes, Yu et al. (31) utilized a machine learning model integrating multimodal images, including DWI, ADC, FLAIR, SWI, and T1w. This model achieved notable performance metrics, with an accuracy of 0.831, sensitivity of 0.739, specificity of 0.902, an F1 score of 0.788, and an AUC of 0.902. Radiomic features extracted using the LightGBM model from

multimodal MRI effectively forecast stroke prognosis, demonstrating the model's high predictive value for clinical outcomes in acute stroke patients. This underscores the potential of multimodal imaging in making precise prognosis predictions for ischemic stroke.

While the feasibility of using machine learning based on radiomics alongside clinical factors for predicting acute ischemic stroke outcomes is recognized, it is crucial to be mindful of possible biases in participant selection across studies. Zhou et al. (32) predicted acute ischemic stroke outcomes by blending radiomic features from multimodal imaging with clinical factors. Their clinical-radiomic nomogram showed superior ROC AUCs in both training and validation groups, achieving 0.868 and 0.890, respectively, surpassing models based solely on clinical or radiomic data. Despite its utility, this study's single-center nature and restriction to DWI and ADC sequences without considering other modalities may introduce biases.

Future research should aim for broader and more diverse datasets, implement stringent study methodologies, and address potential biases to solidify the reliability and applicability of predictive models. Including data from multiple centers and varied patient demographics will strengthen the external validity of research outcomes, paving the way for advancements in personalized stroke management.

Diverging from prior research, our study utilized extensive datasets and integrated radiomic features from multimodal MRI (T1w, T2w, DWI) with clinical risk factors to develop a clinical-radiomic model to forecast the prognosis of acute ischemic stroke. This model exhibited enhanced performance in the validation set, outperforming individual imaging features or clinical factors in discriminative capacity, calibration, and clinical applicability. The model's performance metrics in the training set were noteworthy: sensitivity reached 0.951, specificity 0.811, and accuracy 0.903. Furthermore, in the validation set, sensitivity was 0.918, specificity 0.681, and accuracy 0.825. Despite some variability in these metrics, our model consistently offers valuable insights for clinicians in making informed decisions regarding the treatment and prognosis of acute ischemic stroke.

This research leveraged a broad spectrum of imaging modalities for feature extraction, including T1w, T2w, and both raw and processed DWI images. Twenty-one radiomic features were extracted, covering dimensions such as shape, energy, texture, and various gray-level matrices (GLCM, GLRLM, GLSZM, GLDM). These features provide a comprehensive portrayal of ischemic stroke heterogeneity, offering a sophisticated understanding of stroke pathology. Notably, elevated values in these feature analyses correlate with poorer patient outcomes, enabling precise and quantifiable assessments of imaging characteristics linked to ischemic stroke prognosis.

It's worth noting that our study revealed multivariate logistic regression analysis revealed coronary heart disease, uric acid levels, blood glucose levels, homocysteine, and the platelet-to-lymphocyte ratio (PLR) as independent risk factors affecting stroke prognosis, which align with clinical realities.

Extensive research supports that high blood glucose levels are an independent risk factor for stroke, increasing susceptibility to ischemic stroke by 2–4 times (33). The strong correlation between type 2 diabetes and stroke risk is well-documented across various studies (33, 34). Effective blood glucose management is crucial for reducing the risk of diabetes and its complications, notably the increased risk of stroke.

Coronary heart disease and ischemic stroke often overlap in their underlying mechanisms, with individuals with coronary heart disease at higher risk of cardiovascular events after a stroke. This interaction can lead to complex multi-organ impairment and hinder optimal recovery. A comprehensive treatment approach is essential to manage these interconnected conditions and improve patient outcomes (35).

Uric acid, a product of purine breakdown, has dual roles: it scavenges free radicals and promotes neuronal glutathione synthesis, offering neuroprotection. Elevated uric acid levels are linked to stroke risk due to its oxidative properties, though it also has neuroprotective effects. Maintaining the balance of uric acid is crucial to managing stroke risk (36).

Elevated homocysteine levels, often associated with deficiencies in folate, vitamin B6, and B12, significantly increase stroke risk. For every 5 umol/L rise in homocysteine levels, the risk of stroke escalates by 95%. Higher levels are also linked to early neurological deterioration and increased risk of stroke recurrence and mortality. Monitoring homocysteine levels is crucial in mitigating stroke risk (37).

The platelet-to-lymphocyte ratio (PLR), derived from platelet and lymphocyte counts, serves as an inflammation marker and prognostic indicator for disease progression (38). PLR has predictive value in various conditions, including cancer. In acute ischemic stroke (AIS), heightened PLR levels correlate with larger infarcts and poorer prognosis. PLR also predicts clinical outcomes at 90 days, making it a valuable prognostic biomarker in AIS scenarios (39).

5 Limitation

While this research offers promising insights, it is imperative to recognize its limitations. The study's retrospective design introduces the potential for selection bias and the influence of confounding variables. Furthermore, the study is constrained by a relatively small participant pool and a limited scope for external validation. Future research would benefit from prospective, multicenter studies encompassing larger cohorts and broader validation efforts to bolster the findings' robustness. Additionally, this investigation's treatment of ischemic stroke etiology, especially concerning cerebral infarction locations, lacks granularity. The reliance on manual delineation for regions of interest (ROI) could introduce subjectivity; thus, subsequent studies might enhance accuracy by adopting advanced software solutions, refining training protocols, and minimizing subjective bias. The inconsistency in standard head MRI sequences across various institutions and divergent institutional protocols posed challenges in evaluating a comprehensive range of MRI sequences. Future endeavors should focus on improving patient engagement, augmenting the collection of multimodal MRI data (e.g., FLAIR, ADC, SWAN), and constructing sophisticated models to overcome these limitations. Importantly, the initial study cohort was composed exclusively of patients who did not receive ischemic reperfusion therapy. Comparative analyses of patients undergoing ischemic reperfusion therapy vs. those who do not could provide valuable insights into the predictive utility of radiological markers and clinical indicators in treated acute ischemic stroke patients.

6 Conclusion

In conclusion, this study demonstrates that the incorporation of radiomic features from T1-weighted (T1w), T2-weighted (T2w), and

diffusion-weighted imaging (DWI) with clinical parameters into a cohesive clinical-radiomics model significantly improves the predictive accuracy for the prognosis and treatment responses in ischemic stroke. This advancement fosters a deeper comprehension of therapeutic impacts, prognostic evaluations, and clinical assessments, offering invaluable insights for medical professionals in diagnosis, therapeutic intervention, and rehabilitation. The synergy between radiomic attributes and clinical information heralds a promising avenue for enhancing personalized medicine strategies and elevating patient care outcomes in ischemic stroke.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board at Xi'an Central Hospital and Tongchuan Mining Bureau Hospital. Written informed consent was not required due to the retrospective nature of the study.

Author contributions

KG: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Formal analysis, Project administration, Resources, Software, Validation, Visualization. BZ: Data curation, Writing – review & editing. RL: Data curation, Investigation, Writing – review & editing. JX: Data curation, Investigation, Writing – review & editing. QW: Methodology, Software, Visualization, Writing – review & editing. KC: Data curation, Investigation, Writing – review & editing. YS: Data curation, Investigation, Writing – review & editing. JL: Investigation, Writing – review & editing, Project administration, Supervision. WC: Supervision, Writing – review & editing, Project administration. ZL: Project administration, Supervision, Writing – review & editing. ZD: Writing – review & editing, Funding acquisition. NG: Writing – original draft, Writing – review & editing, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study received financial support from the Natural Science Foundation of Shaanxi Province (Project Code: 2021JM562), the Shaanxi Provincial Science and Technology Plan Project (2022SF-418), the Science and Technology Planning Project of Xi'an City (Project Code: 2023YXYJ0036), and the Xi'an Central Hospital Scientific Research Fund (2024YB04).

Acknowledgments

The authors express our gratitude to all individuals involved in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be perceived as potential conflicts of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet*. (2020) 396:1204–22. doi: 10.1016/S0140-6736(20)30925-9

2. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. 333:1581–7. doi: 10.1056/NEJM199512143332401

3. Hacke W, Kaste M, Bluhmki E, Brozman M, Dávalos A, Guidetti D, et al. Thrombolysis with Alteplase 3 to 4.5 hours after acute ischemic stroke. *N Engl J Med Overseas Ed*. 359:1317–29. doi: 10.1056/NEJMoa0804656

4. Bracard S, Ducrocq X, Mas JL, Soudant M, Oppenheim C, Moulin T, et al. Investigators T: mechanical thrombectomy after intravenous alteplase versus alteplase alone after stroke (THRACE): a randomised controlled trial. *Lancet Neurol*. (2016) 15:1138–47. doi: 10.1016/S1474-4422(16)30177-6

5. Yao M, Ren Y, Jia Y, Xu J, Wang Y, Zou K, et al. Projected burden of stroke in China through 2050. *Chin Med J*. (2023) 136:1598–605. doi: 10.1097/CM9.0000000000002060

6. Wang YJ, Li ZX, Gu HQ, Zhai Y, Jiang Y, Zhao XQ, et al. China stroke statistics 2019: a report from the National Center for healthcare quality Management in Neurological Diseases, China National Clinical Research Center for neurological Diseases, the Chinese Stroke Association, National Center for chronic and non-communicable disease control and prevention, Chinese Center for Disease Control and Prevention and Institute for global neuroscience and stroke collaborations. *Stroke Vasc Neurol*. (2020) 5:211–39. doi: 10.1136/svn-2020-000457

7. Eslami V, Tahsili-Fahadan P, Rivera-Lara L, Gandhi D, Ali H, Parry-Jones A, et al. Influence of intracerebral hemorrhage location on outcomes in patients with severe intraventricular hemorrhage. *Stroke*. (2019) 50:1688–95. doi: 10.1161/STROKEAHA.118.024187

8. Johnston KC, Wagner DP, Haley EC Jr, Connors AF Jr. Stroke RIRToTMiA: combined clinical and imaging information as an early stroke outcome measure. *Stroke*. (2002) 33:466–72. doi: 10.1161/hs0202.102881

9. Zhao J, Feng J, Ma Q, Li C, Qiu F. Prognostic value of inflammation biomarkers for 30-day mortality in critically ill patients with stroke. *Front Neurol*. (2023) 14:1110347. doi: 10.3389/fneur.2023.1110347

10. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging-"how-to" guide and critical reflection. *Insights Imaging*. (2020) 11:91. doi: 10.1186/s13244-020-00887-2

11. Rogers W, Thulasi Seetha S, Refaee TAG, Lieverse RIY, Granzier RWY, Ibrahim A, et al. Radiomics: from qualitative to quantitative imaging. *Br J Radiol*. (2020) 93:20190948. doi: 10.1259/bjr.20190948

12. Yang Y, Tang L, Deng Y, Li X, Luo A, Zhang Z, et al. The predictive performance of artificial intelligence on the outcome of stroke: a systematic review and meta-analysis. *Front Neurosci*. (2023) 17:1256592. doi: 10.3389/fnins.2023.1256592

13. Chen Q, Xia T, Zhang M, Xia N, Liu J, Yang Y. Radiomics in stroke neuroimaging: techniques, applications, and challenges. *Aging Dis*. (2021) 12:143–54. doi: 10.14336/AD.2020.0421

14. Quan G, Ban R, Ren JL, Liu Y, Wang W, Dai S, et al. FLAIR and ADC image-based Radiomics features as predictive biomarkers of unfavorable outcome in patients with acute ischemic stroke. *Front Neurosci*. (2021) 15:730879. doi: 10.3389/fnins.2021.730879

15. Zhang Y, Zhuang Y, Ge Y, Wu PY, Zhao J, Wang H, et al. MRI whole-lesion texture analysis on ADC maps for the prognostic assessment of ischemic stroke. *BMC Med Imaging*. (2022) 22:115. doi: 10.1186/s12880-022-00845-y

16. Wang L, Liu Z, Liang R, Wang W, Zhu R, Li J, et al. Comprehensive. Machine-learning survival framework develops a consensus model in large-scale multicenter cohorts for pancreatic cancer. *eLife*. (2022) 11:150. doi: 10.7554/eLife.80150

17. Alexandra Z, Dongyu L, Rhema V, Kalyan V: Sibyl: understanding and addressing the. Usability challenges of machine learning in high-stakes decision making. IEEE Transactions on Visualization and Computer Graphics. arXiv [Preprint]. arXiv:2103.02071v2. (2021).

18. Guo Y, Yang Y, Cao F, Li W, Wang M, Luo Y, et al. Novel. Survival features generated by clinical text information and Radiomics features may improve the prediction of ischemic stroke outcome. *Diagnostics*. (2022) 12:12(7). doi: 10.3390/diagnostics12071664

19. Barber PA, Demchuk AM, Zhang J, Buchan AM. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS study group. Alberta stroke Programme early CT score. *Lancet*. (2000) 355:1670–4. doi: 10.1016/S0140-6736(00)02237-6

20. Pop NO, Tit DM, Diaconu CC, Munteanu MA, Babes EE, Stoicescu M, et al. The Alberta stroke program early CT score (ASPECTS): a predictor of mortality in acute ischemic stroke. *Exp Ther Med*. (2021) 22:1371. doi: 10.3892/etm.2021.10805

21. Chen Z, Shi Z, Lu F, Li L, Li M, Wang S, et al. Validation of two automated ASPECTS software on non-contrast computed tomography scans of patients with acute ischemic stroke. *Front Neurol*. (2023) 14:1170955. doi: 10.3389/fneur.2023.1170955

22. Kuang H, Menon BK, Sohn SI, Qiu W. EIS-net: segmenting early infarct and scoring ASPECTS simultaneously on non-contrast CT of patients with acute ischemic stroke. *Med Image Anal*. (2021) 70:101984. doi: 10.1016/j.media.2021.101984

23. Naganuma M, Tachibana A, Fuchigami T, Akahori S, Okumura S, Yi K, et al. Alberta stroke program early CT score calculation using the deep learning-based brain hemisphere comparison algorithm. *J Stroke Cerebrovasc Dis*. (2021) 30:105791. doi: 10.1016/j.jstrokecerebrovasdis.2021.105791

24. Qian Q, Huang HT, Xu L, Jin P, Lin M. Prediction of infarct lesion volumes by processing magnetic resonance apparent diffusion coefficient maps in patients with acute ischemic stroke. *J Stroke Cerebrovasc Dis*. (2016) 25:2821–7. doi: 10.1016/j.jstrokecerebrovasdis.2016.07.041

25. Ma L, Gao PY, Hu QM, Lin Y, Jing LN, Xue J, et al. Effect of baseline magnetic resonance imaging (MRI) apparent diffusion coefficient lesion volume on functional outcome in ischemic stroke. *Neurol Res*. (2011) 33:494–502. doi: 10.1179/016164111X13007856084124

26. Nicolae O, Pop DCZ, Pantiş C, Mekeres F. Clinicopathological evaluation of Moyamoya disease. Case report and review of literature. *Roman J Milit Med*. (2020) CXXIII:5. doi: 10.55453/rjmm.2020.123.2.5

27. Kim T, Heo J, Jang DK, Sunwoo L, Kim J, Lee KJ, et al. Machine learning for detecting moyamoya disease in plain skull radiography using a convolutional neural network. *EBioMedicine*. (2019) 40:636–42. doi: 10.1016/j.ebiom.2018.12.043

28. Qin K, Guo Z, Peng C, Gan W, Zhou D, Chen G. Prediction of the mean transit time using machine learning models based on radiomics features from digital subtraction angiography in moyamoya disease or moyamoya syndrome-a development and validation model study. *Cardiovasc Diagn Ther*. (2023) 13:879–92. doi: 10.21037/cdt-23-151

29. Dragos HM, Stan A, Pintican R, Feier D, Lebovici A, Panaitescu PS, et al. MRI Radiomics and predictive models in assessing ischemic stroke outcome-a systematic review. *Diagnostics*. (2023) 13:857. doi: 10.3390/diagnostics13050857

30. Wang H, Sun Y, Ge Y, Wu PY, Lin J, Zhao J, et al. A clinical-Radiomics nomogram for functional outcome predictions in ischemic stroke. *Neurol Ther*. (2021) 10:819–32. doi: 10.1007/s40120-021-00263-2

31. Yu H, Wang Z, Sun Y, Bo W, Duan K, Song C, et al. Prognosis of ischemic stroke predicted by machine learning based on multi-modal MRI radiomics. *Front Psych*. (2022) 13:1105496. doi: 10.3389/fpsyt.2022.1105496

32. Zhou Y, Wu D, Yan S, Xie Y, Zhang S, Lv W, et al. Feasibility of a clinical-Radiomics model to predict the outcomes of acute ischemic stroke. *Korean J Radiol*. (2022) 23:811–20. doi: 10.3348/kjr.2022.0160

33. The Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet*. (2010) 375:2215–22. doi: 10.1016/S0140-6736(10)60484-9

34. Guo L, Yu M, Zhong J, Wu H, Pan J, Gong W, et al. Stroke risk among patients with type 2 diabetes mellitus in Zhejiang: a population-based prospective study in China. *Int J Endocrinol*. (2016) 2016:1–8. doi: 10.1155/2016/6380620

35. Yousufuddin M, Bartley AC, Alsawas M, Sheely HL, Shultz J, Takahashi PY, et al. Impact of multiple chronic conditions in patients hospitalized with stroke and transient ischemic attack. *J Stroke Cerebrovasc Dis*. (2017) 26:1239–48. doi: 10.1016/j.jstrokecerebrovasdis.2017.01.015

36. Qiao T, Wu H, Peng W. The relationship between elevated serum uric acid and risk of stroke in adult: an updated and dose-response Meta-analysis. *Front Neurol*. (2021) 12:674398. doi: 10.3389/fneur.2021.674398

37. The Homocyteine Studies Collaboation. Homocysteine and risk of ischemic heart Diease and stroke a meta-analysis. *JAMA*. (2015) 288:2015. doi: 10.1001/jama.288.16.2015

38. Altintas O, Altintas MO, Tasal A, Kucukdagli OT, Asil T. The relationship of platelet-to-lymphocyte ratio with clinical outcome and final infarct core in acute ischemic stroke patients who have undergone endovascular therapy. *Neurol Res*. (2016) 38:759–65. doi: 10.1080/01616412.2016.1215030

39. Chen C, Gu L, Chen L, Hu W, Feng X, Qiu F, et al. Neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio as potential predictors of prognosis in acute ischemic stroke. *Front Neurol*. (2021) 11:525621. doi: 10.3389/fneur.2020.525621

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 17 June 2024
DOI 10.3389/fneur.2024.1373306

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Cristiano Capurso,
University of Foggia, Italy
Mengyang Wang,
Capital Medical University, China

*CORRESPONDENCE
Ye Jiang
13383263807@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 25 January 2024
ACCEPTED 06 June 2024
PUBLISHED 17 June 2024

CITATION
Li N, Gao Y, Li L-t, Hu Y-d, Ling L, Jia N, Chen Y-j, Meng Y-n and Jiang Y (2024) Development and validation of a nomogram predictive model for cognitive impairment in cerebral small vessel disease: a comprehensive retrospective analysis. *Front. Neurol.* 15:1373306. doi: 10.3389/fneur.2024.1373306

COPYRIGHT
© 2024 Li, Gao, Li, Hu, Ling, Jia, Chen, Meng and Jiang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development and validation of a nomogram predictive model for cognitive impairment in cerebral small vessel disease: a comprehensive retrospective analysis

Ning Li[1†], Yan Gao[2†], Li-tao Li[3], Ya-dong Hu[1], Li Ling[1], Nan Jia[1], Ya-jing Chen[1], Ya-nan Meng[1] and Ye Jiang[1*]

[1]Department of Neurology, Affiliated Hospital of Hebei University, Baoding, China, [2]Department of Neurology, Renmin Hospital of Wuhan University, Wuhan, China, [3]Department of Neurology, Hebei General Hospital, Shijiazhuang, China

Background: Cerebral small vessel disease (CSVD) is a common neurodegenerative condition in the elderly, closely associated with cognitive impairment. Early identification of individuals with CSVD who are at a higher risk of developing cognitive impairment is crucial for timely intervention and improving patient outcomes.

Objective: The aim of this study is to construct a predictive model utilizing LASSO regression and binary logistic regression, with the objective of precisely forecasting the risk of cognitive impairment in patients with CSVD.

Methods: The study utilized LASSO regression for feature selection and logistic regression for model construction in a cohort of CSVD patients. The model's validity was assessed through calibration curves and decision curve analysis (DCA).

Results: A nomogram was developed to predict cognitive impairment, incorporating hypertension, CSVD burden, apolipoprotein A1 (ApoA1) levels, and age. The model exhibited high accuracy with AUC values of 0.866 and 0.852 for the training and validation sets, respectively. Calibration curves confirmed the model's reliability, and DCA highlighted its clinical utility. The model's sensitivity and specificity were 75.3 and 79.7% for the training set, and 76.9 and 74.0% for the validation set.

Conclusion: This study successfully demonstrates the application of machine learning in developing a reliable predictive model for cognitive impairment in CSVD. The model's high accuracy and robust predictive capability provide a crucial tool for the early detection and intervention of cognitive impairment in patients with CSVD, potentially improving outcomes for this specific condition.

KEYWORDS

cerebral small vessel disease, cognitive impairment, predictive modeling, LASSO regression, logistic regression, nomogram

1 Introduction

Cerebral small vessel disease (CSVD) is intricately linked to cognitive decline and represents a critical focus in the study of vascular contributions to cognitive impairment and dementia (1–3). This spectrum of pathological processes has been increasingly recognized for its substantial impact on public health, especially given the aging global population (2, 4–7). Despite its clinical importance, the detection and quantification of cognitive impairment attributable to CSVD remain fraught with complexities. Cognitive symptoms often manifest subtly and progress insidiously, making early diagnosis a formidable challenge in clinical settings (8).

The traditional approach to diagnosing CSVD-related cognitive impairment relies heavily on clinical judgment, which is subject to variability and may fail to capture the nuanced progression of the disease. Consequently, there is a pressing need for objective and reproducible diagnostic tools that can accurately predict the onset and trajectory of cognitive decline in CSVD. Addressing this need, our study introduces a predictive model that synthesizes demographic, clinical, neuroimaging, and biomarker data to objectively assess the risk of cognitive decline. The current models in literature have either not incorporated such a diverse dataset or have not been validated sufficiently for clinical application (9, 10). Our approach utilizes advanced machine learning techniques, including LASSO regression for feature selection and logistic regression for model development, filling a critical gap by providing a tool with both high sensitivity and specificity for CSVD cognitive sequelae.

This study leverages a comprehensive dataset of 377 CSVD patients, encompassing a wide array of variables including demographic details, clinical history, neuroimaging findings, and biomarkers. Through rigorous statistical methodologies, we aim to construct a predictive model that not only distinguishes between patients with and without cognitive impairment but also provides a probabilistic assessment of the risk of cognitive decline. By adopting machine learning techniques, particularly LASSO regression for feature selection followed by logistic regression for model development, we seek to create a model that is both sensitive and specific to the cognitive sequelae of CSVD.

In doing so, our objectives are twofold: firstly, to present a model that can be readily applied in clinical practice for the early identification of patients at risk of cognitive impairment due to CSVD, and secondly, to contribute to the body of knowledge that underpins the intersection of neurovascular pathology and neurodegeneration. It is our anticipation that such a model will not only facilitate early intervention strategies but also spur further research into targeted therapies for this underdiagnosed yet prevalent condition.

2 Materials and methods

2.1 Study population and design

In a cross-sectional study conducted from July 2022 to October 2023 at the Neurology Department of the Affiliated Hospital of Hebei University, we targeted a cohort of patients aged 50 and above. These participants were subjected to comprehensive cranial magnetic resonance imaging (MRI) to confirm CSVD diagnosis. The inclusion criteria mandated completion of a full cranial MRI series, provision of serum samples for ELISA testing to assess inflammatory markers, and undergoing the Montreal Cognitive Assessment (MoCA) to gauge cognitive function. The MoCA, a widely acknowledged tool for detecting mild cognitive impairments, was specifically chosen for its robust validity across diverse age groups and clinical conditions (11, 12), thereby serving as an optimal instrument for assessing CSVD-related cognitive impairments. The scoring of MoCA was conducted by trained personnel, adhering to standardized procedures to ensure consistency and reliability of cognitive function assessment. Exclusion criteria encompassed the presence of neurological conditions other than CSVD that might impair cognitive function, a history of psychiatric disorders or current use of psychoactive drugs, contraindications to MRI such as claustrophobia or presence of metal implants, inadequate quality of cranial MRI for reliable assessment, acute infections and severe systemic illnesses that could impede participation in the study. Participants' cognitive status was categorized based on MoCA scores, with scores of 25 or below indicating cognitive impairment, and scores of 26 or above denoting normal cognitive function. Alongside MRI diagnostics, participants underwent routine lab tests including complete blood count, renal and liver function tests, and electrolyte panels. Informed consent for serum sample collection for ELISA testing was obtained in line with ethical standards. This study was conducted in strict adherence to ethical guidelines and was approved by the Ethics Committee of the Affiliated Hospital of Hebei University, with the approval number HDFYLL-KY-2023-060. This approval ensures that our study complies with the ethical principles outlined in the 1964 Helsinki Declaration and its subsequent amendments, reaffirming our commitment to the highest standards of research ethics.

2.2 MRI acquisition and assessment in CSVD patients

In our study, all participants were subjected to brain MRI examinations using a state-of-the-art 3.0 T GE scanner. The identification of CSVD was reliant on the detection of specific neuroimaging indicators, which included any combination of lacunes, white matter hyperintensities (WMH), enlarged perivascular spaces (EPVS), and cerebral microbleeds (CMBs). In this context, WMH were defined as regions showing elevated signal intensity on T2-weighted images, often symmetrically distributed across brain hemispheres. Lacunes were described as round or ovoid cerebrospinal fluid-signal lesions with diameters ranging between 3 and 15 mm. Additionally, CMBs manifested as small, round, hypointense areas, 2–10 mm in size, evident in susceptibility-weighted imaging sequences. In line with the comprehensive CSVD scoring methodology initially developed by Wardlaw et al. (5, 11), we evaluated the total CSVD burden employing an ordinal scale that spans from 0 to 4. This evaluation process included assigning one point for each of the following criteria: the severity of WMH, as indicated by a periventricular WMH score of 3 or a deep WMH score between 2 and 3; the presence of lacunes; the existence of CMBs; and the significant presence of basal ganglia EPVS, particularly exceeding a count of 10 (12). For an accurate and unbiased evaluation of these CSVD markers, neuroimaging specialists Yan Hou and Huan Zhou, devoid of access to the participants' clinical information, undertook the assessment. Their analysis adhered rigorously to the Standards for

Reporting Vascular Changes on Neuroimaging (STRIVE) criteria (11, 13), which provided a framework for consistent and reliable interpretation of CSVD-related neuroimaging findings.

2.3 Clinical blood biochemistry assessment

In this study, we performed an extensive assessment of clinical blood biochemistry parameters in patients with CSVD. Our evaluation included a broad spectrum of 44 laboratory markers, categorically divided into routine and specialized tests. Routine assessments comprised complete blood count, renal and liver function tests, electrolyte balance, lipid profiles, and coagulation parameters. In addition to these standard measures, we conducted a detailed analysis of inflammatory markers, crucial in elucidating CSVD's pathophysiological landscape. These markers encompassed cytokines, acute phase reactants, and specific inflammatory indices, offering insights into the inflammatory status of CSVD patients. Notably, all laboratory data were transformed into binary categorical variables, based on either their median values or clinically established cutoffs. This transformation enabled a nuanced exploration of the potential correlations between these biochemical markers and cognitive impairment in CSVD, forming a critical component of our predictive model development.

2.4 Clinical evaluation

In this study, we conducted a thorough clinical evaluation of 377 patients diagnosed with CSVD. The participants' demographic information, including age and sex, was meticulously recorded. Medical histories were detailed, focusing on the presence of hypertension, diabetes, and hypercholesterolemia. Hypertension was defined as having a systolic blood pressure ≥140 mmHg, diastolic blood pressure ≥90 mmHg, or being under antihypertensive medication. Diabetes was identified by fasting blood glucose levels ≥7.0 mmol/L, OGTT2h levels ≥11.1 mmol/L, or the use of hypoglycemic medications. Hypercholesterolemia was recognized by total cholesterol or LDL cholesterol levels exceeding normal range limits. Additionally, lifestyle factors such as smoking and alcohol consumption histories were gathered, alongside information on the duration of the disease.

2.5 Statistical methodology

In our investigation involving 377 CSVD patients, the cohort was randomly divided into a training set with 265 participants and a validation set comprising 112 participants, following a 7:3 allocation ratio. This randomization ensured a balanced and unbiased approach to model development. we opted for the stratification of continuous laboratory variables into tertiles. This approach was employed to facilitate a more meaningful interpretation of the data by categorizing it into low, medium, and high ranges. Utilizing tertiles allows for a clear delineation of patient subgroups based on their laboratory values, aligning with the clinical relevance of these categories. The selection of thresholds for continuous variables was informed by clinical expertise and mirrored established norms in contemporary research and statistical methodologies. For categorical data, we presented frequencies and percentages. To ascertain baseline similarities or disparities between the training and validation cohorts, we utilized appropriate statistical tests, opting for the χ^2 test or Fisher's exact test for categorical data.

Within the training dataset, the LASSO regression method was employed to pinpoint crucial risk factors linked to cognitive impairment in CSVD patients. LASSO regression was utilized for its efficacy in both variable selection and regularization, enabling the identification of the most predictive features for cognitive impairment in CSVD while mitigating overfitting. This technique helped in selecting variables with non-zero coefficients for further analysis. These identified variables were then incorporated into a logistic regression model to ascertain independent predictors of cognitive impairment in CSVD. We constructed a nomogram from these identified risk factors, derived from the logistic regression model. The nomogram's predictive performance was evaluated using the receiver operating characteristic curve (AUC-ROC), and calibration curves were generated to align predicted probabilities with actual outcomes. Furthermore, the clinical utility of the model was assessed using decision curve analysis (DCA). To mitigate overfitting, we employed several strategies. First, we used LASSO regression to perform variable selection and reduce model complexity. Second, we conducted 10-fold cross-validation to determine the optimal regularization parameter (λ), ensuring the model's performance on multiple data subsets. Finally, we validated the model using an independent dataset to test its generalizability. These combined strategies enhance the model's robustness and reduce the risk of overfitting. All statistical procedures were carried out using R software (version 4.3.0),[1] and a p-value of less than 0.05 (two-tailed) was considered to denote statistical significance.

3 Results

3.1 Baseline characteristics

From July 2022 to October 2023, this investigation initially recruited 427 participants who satisfied the inclusion benchmarks. Subsequent application of exclusion criteria led to the withdrawal of 50 participants, yielding a dataset of 377 individuals for subsequent analyses as depicted in Figure 1. Of these, 265 formed the training cohort, while the remaining 112 were allocated to the validation cohort. Table 1 delineates the demographic and clinical profiles at baseline for both cohorts. This study incorporated the following 52 potential indicators related to cognitive impairment in CSVD: CSVD Burden, Gender, Smoking, Drinking, Hypertension, Diabetes, Hyperlipidemia, Age, White Blood Cell Count, Platelet Count, Neutrophil Count, Lymphocyte Count, Monocyte Count, Urea, Creatinine, Uric Acid, Prothrombin Time, Activated Partial Thromboplastin Time, Thrombin Time, Fibrinogen, Hemoglobin A1c, Total Cholesterol, Triglycerides, High-Density Lipoprotein, Low-Density Lipoprotein, Very Low-Density Lipoprotein, Apolipoprotein A1 (ApoA1), Apolipoprotein B100, ApoA1 to ApoB100 Ratio, Lipoprotein(a), Total Protein, Albumin, Globulin,

1 http://www.r-project.org/

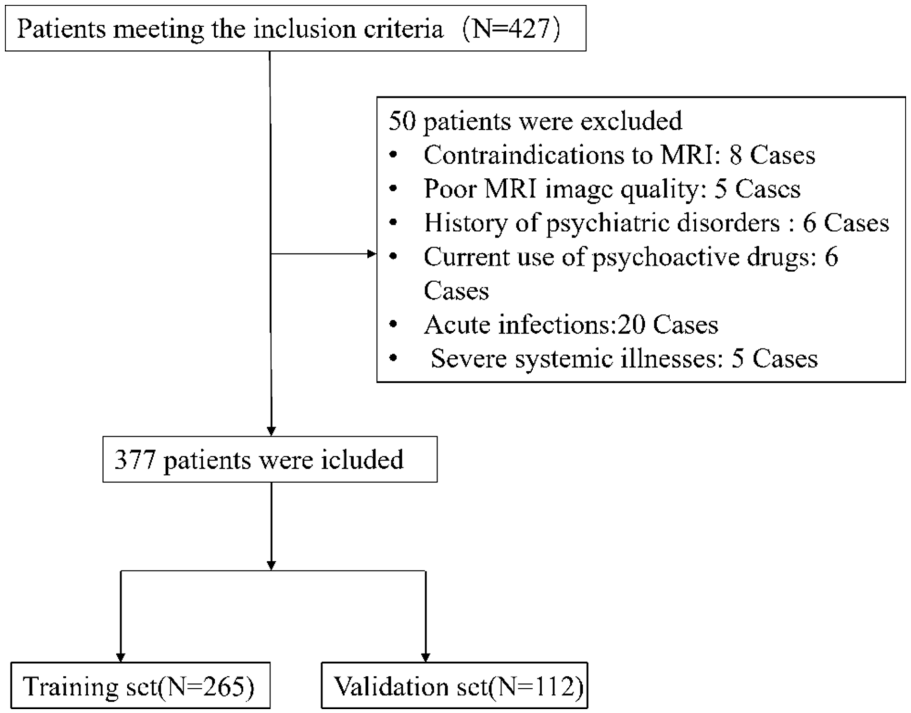


FIGURE 1
Patient selection flowchart. This diagram details the process of selecting eligible CSVD patients for the study, highlighting the inclusion and exclusion criteria. It shows the initial number of participants, the application of exclusion criteria, and the final number of patients included in the analysis.

Albumin to Globulin Ratio, Homocysteine, Systemic Immune-Inflammatory Index, Systemic Inflammation Response Index, Neutrophil-to-Lymphocyte Ratio, Lymphocyte-to-Monocyte Ratio, Neutrophil-to-Monocyte Ratio, Neutrophil-to-HDL Ratio, Lymphocyte-to-HDL Ratio, Monocyte-to-HDL Ratio, IL-6, TNFα, VCAM-1, LP-PLA2, CD40L, E-Selectin, ADMA, vWF, and ICAM-1. The lack of significant disparities in the 52 evaluated variables between the training and validation sets underscores the homogeneity of the study population. This parity is critical as it suggests that the predictive model derived from the training set has the potential for generalizability to other patient cohorts, affirming the robustness of the model's predictive capacity for cognitive impairment associated with CSVD (Table 1).

3.2 Variable selection

In our endeavor to elucidate the variables significantly associated with cognitive impairment in CSVD, we rigorously analyzed a dataset comprising 52 variables, which spanned demographic information, clinical history, and a wide array of laboratory measurements. Employing LASSO regression, a method recognized for its efficiency in variable selection and overfitting prevention, we utilized the glmnet package in R, integrating a 10-fold cross-validation technique to determine the optimal regularization parameter (λ). The selection of λ was guided by the one standard error rule from the minimum criterion in cross-validation error, ensuring the model's parsimony without compromising its predictive accuracy (Figures 2A,B).

This meticulous process distilled the multitude of factors down to 4 pivotal indicators that bear a statistically significant relationship with cognitive impairment in CSVD patients. These indicators-Hypertension, CSVD Burden, ApoA1, and Age-demonstrated the strongest associations with cognitive impairment, underscoring their importance in the predictive model (Table 2). During the LASSO regression process, coefficients of less important variables shrink towards zero. We selected features with non-zero coefficients, which ensures that only the most relevant predictors are included in the final model. This approach balances model simplicity and predictive performance, enhancing both interpretability and robustness of the model. Although the coefficient for CSVD Burden is relatively small, it was retained due to its clinical significance in reflecting the overall severity of cerebral small vessel disease.

3.3 Multivariable analysis

Our multivariable logistic regression analysis, adjusting for potential confounders, identified significant associations between CSVD-related cognitive impairment and several key factors from the 12 variables initially pinpointed by LASSO regression. Notably, Hypertension (OR: 1.78, 95% CI: 1.32–3.84, $p < 0.001$) and Age (OR: 1.85, 95% CI: 1.55–3.55, $p < 0.001$) were potent predictors of impairment, indicating substantially elevated risks. A higher CSVD Burden also heightened the risk (OR: 1.83, 95% CI: 1.23–2.72, $p = 0.003$), whereas increased ApoA1 levels demonstrated a protective effect (OR: 0.64, 95% CI: 0.43–0.97, $p = 0.034$). These variables reflect critical demographic and biological aspects influencing CSVD prognosis, as detailed in Table 3.

TABLE 1 Comparative analysis of potential predictive factors in a cerebral small vessel disease (CSVD) cognitive impairment predictive model between training and validation sets.

Variables	Total (*n* = 377)	training set (*n* = 265)	validation set (*n* = 112)	*p*-value
Cognitive impairment, *n* (%)				0.438
No	132 (35)	89 (34)	43 (38)	
Yes	245 (65)	176 (66)	69 (62)	
CSVD.Burden, *n* (%)				0.912
0–1	173 (46)	120 (45)	53 (47)	
2	89 (24)	64 (24)	25 (22)	
3–4	115 (31)	81 (31)	34 (30)	
Gender, *n* (%)				0.273
Female	183 (49)	134 (51)	49 (44)	
Male	194 (51)	131 (49)	63 (56)	
Smoking, *n* (%)				0.8
No	261 (69)	185 (70)	76 (68)	
Yes	116 (31)	80 (30)	36 (32)	
Drinking, *n* (%)				0.174
No	124 (33)	81 (31)	43 (38)	
Yes	253 (67)	184 (69)	69 (62)	
Hypertension, *n* (%)				0.3
No	103 (27)	77 (29)	26 (23)	
Yes	274 (73)	188 (71)	86 (77)	
Diabetes, *n* (%)				0.3
No	267 (71)	183 (69)	84 (75)	
Yes	110 (29)	82 (31)	28 (25)	
Hyperlipidemia, *n* (%)				0.956
No	250 (66)	175 (66)	75 (67)	
Yes	127 (34)	90 (34)	37 (33)	
Age, *n* (%)				0.9
≤69	249 (66)	174 (66)	75 (67)	
>69	128 (34)	91 (34)	37 (33)	
White blood cell count ($\times10^9$/L), *n* (%)				0.842
<6.04	126 (33)	91 (34)	35 (31)	
6.04–7.70	128 (34)	89 (34)	39 (35)	
>7.70	123 (33)	85 (32)	38 (34)	
Platelet count ($\times10^9$/L), *n* (%)				0.973
<200	126 (33)	89 (34)	37 (33)	
200–250	128 (34)	89 (34)	39 (35)	
>250	123 (33)	87 (33)	36 (32)	
Neutrophil count ($\times10^9$/L), *n* (%)				0.581
<3.83	127 (34)	86 (32)	41 (37)	
3.83–5.34	125 (33)	92 (35)	33 (29)	
>5.34	125 (33)	87 (33)	38 (34)	
Lymphocyte count ($\times10^9$/L), *n* (%)				0.533
<130	126 (33)	93 (35)	33 (29)	
1.30–1.74	126 (33)	85 (32)	41 (37)	
>1.74	125 (33)	87 (33)	38 (34)	

(Continued)

TABLE 1 (Continued)

Variables	Total (*n* = 377)	training set (*n* = 265)	validation set (*n* = 112)	*p*-value
Monocyte count (×10^9/L), *n* (%)				0.612
<0.40	128 (34)	87 (33)	41 (37)	
0.40–0.54	128 (34)	94 (35)	34 (30)	
>0.54	121 (32)	84 (32)	37 (33)	
Urea (mmol/L), *n* (%)				0.791
<5.00	120 (32)	87 (33)	33 (29)	
5.00–6.19	120 (32)	84 (32)	36 (32)	
>6.19	137 (36)	94 (35)	43 (38)	
Creatinine (μmol/L), *n* (%)				0.482
<57	128 (34)	88 (33)	40 (36)	
57–70	128 (34)	95 (36)	33 (29)	
>70	121 (32)	82 (31)	39 (35)	
Uric acid (μmol/L), *n* (%)				0.301
<260	127 (34)	89 (34)	38 (34)	
260–325	127 (34)	95 (36)	32 (29)	
>325	123 (33)	81 (31)	42 (38)	
Prothrombin time (s), *n* (%)				0.308
<11.1	143 (38)	94 (35)	49 (44)	
11.1–11.6	111 (29)	82 (31)	29 (26)	
>11.6	123 (33)	89 (34)	34 (30)	
Activated partial thromboplastin time (s), *n* (%)				0.273
<27.6	126 (33)	82 (31)	44 (39)	
27.6–31.9	128 (34)	92 (35)	36 (32)	
>31.9	123 (33)	91 (34)	32 (29)	
Thrombin time (s), *n* (%)				0.305
<14.5	131 (35)	87 (33)	44 (39)	
14.5–17.4	121 (32)	91 (34)	30 (27)	
>17.4	125 (33)	87 (33)	38 (34)	
Fibrinogen (g/L), *n* (%)				0.954
<2.68	130 (34)	91 (34)	39 (35)	
2.68–3.26	122 (32)	87 (33)	35 (31)	
>3.26	125 (33)	87 (33)	38 (34)	
Hemoglobin A1c (mmol/L), *n* (%)				0.423
<5.6	132 (35)	95 (36)	37 (33)	
5.6–6.2	133 (35)	88 (33)	45 (40)	
>6.2	112 (30)	82 (31)	30 (27)	
Total cholesterol (mmol/L), *n* (%)				0.874
<3.92	128 (34)	91 (34)	37 (33)	
3.92–4.97	124 (33)	85 (32)	39 (35)	
>4.97	125 (33)	89 (34)	36 (32)	
Triglycerides (mmol/L), n (%)				0.24
<1.00	128 (34)	88 (33)	40 (36)	
1.00–1.47	125 (33)	83 (31)	42 (38)	
>1.47	124 (33)	94 (35)	30 (27)	
High-density lipoprotein (mmol/L), n (%)				0.124

(Continued)

TABLE 1 (Continued)

Variables	Total (*n* = 377)	training set (*n* = 265)	validation set (*n* = 112)	*p*-value
<1.02	129 (34)	96 (36)	33 (29)	
1.02–1.23	128 (34)	93 (35)	35 (31)	
>1.23	120 (32)	76 (29)	44 (39)	
Low-density lipoprotein (mmol/L), *n* (%)				0.707
<2.39	124 (33)	86 (32)	38 (34)	
2.39–3.23	127 (34)	87 (33)	40 (36)	
>3.23	126 (33)	92 (35)	34 (30)	
Very low-density lipoprotein (mmol/L), *n* (%)				0.983
<0.36	127 (34)	90 (34)	37 (33)	
0.36–0.54	126 (33)	88 (33)	38 (34)	
>0.54	124 (33)	87 (33)	37 (33)	
Apolipoprotein A1 (g/L), *n* (%)				0.083
<0 0.97	127 (34)	91 (34)	36 (32)	
0.97–1.12	128 (34)	97 (37)	31 (28)	
>1.12	122 (32)	77 (29)	45 (40)	
Apolipoprotein B100 (g/L), *n* (%)				0.982
<0.68	130 (34)	92 (35)	38 (34)	
0.68–0.88	125 (33)	88 (33)	37 (33)	
>0.88	122 (32)	85 (32)	37 (33)	
ApoA1 to ApoB100 ratio, *n* (%)				0.394
<1.19	126 (33)	94 (35)	32 (29)	
1.19–1.55	127 (34)	85 (32)	42 (38)	
>1.55	124 (33)	86 (32)	38 (34)	
Lipoprotein(a) (mg/L), *n* (%)				0.675
<121	126 (33)	89 (34)	37 (33)	
121–678	140 (37)	95 (36)	45 (40)	
>678	111 (29)	81 (31)	30 (27)	
Total protein (g/L), *n* (%)				0.361
<62	143 (38)	95 (36)	48 (43)	
62–67	125 (33)	93 (35)	32 (29)	
>67	109 (29)	77 (29)	32 (29)	
Albumin (g/L), *n* (%)				0.424
<35	159 (42)	110 (42)	49 (44)	
35–40	124 (33)	84 (32)	40 (36)	
>40	94 (25)	71 (27)	23 (21)	
Globulin (g/L), *n* (%)				0.272
<26	165 (44)	109 (41)	56 (50)	
26–28	106 (28)	77 (29)	29 (26)	
>28	106 (28)	79 (30)	27 (24)	
Albumin to globulin ratio, *n* (%)				0.653
<1.38	134 (36)	98 (37)	36 (32)	
1.38–1.56	119 (32)	81 (31)	38 (34)	
>1.56	124 (33)	86 (32)	38 (34)	
Homocysteine (μmol/L), *n* (%)				0.505
<15	135 (36)	94 (35)	41 (37)	

(Continued)

TABLE 1 (Continued)

Variables	Total (*n* = 377)	training set (*n* = 265)	validation set (*n* = 112)	*p*-value
15–20	123 (33)	91 (34)	32 (29)	
>20	119 (32)	80 (30)	39 (35)	
Systemic immune-inflammatory index, *n* (%)				0.928
<515.27	126 (33)	89 (34)	37 (33)	
515.27–878.63	126 (33)	87 (33)	39 (35)	
>878.63	125 (33)	89 (34)	36 (32)	
Systemic inflammation response index, *n* (%)				0.761
<1.02	126 (33)	89 (34)	37 (33)	
1.02–1.75	126 (33)	91 (34)	35 (31)	
>1.75	125 (33)	85 (32)	40 (36)	
Neutrophil-to-lymphocyte ratio, *n* (%)				0.918
<2.36	126 (33)	87 (33)	39 (35)	
2.36–3.64	126 (33)	90 (34)	36 (32)	
>3.64	125 (33)	88 (33)	37 (33)	
Lymphocyte-to-monocyte ratio, *n* (%)				0.627
<2.86	126 (33)	92 (35)	34 (30)	
2.86–3.93	126 (33)	85 (32)	41 (37)	
>3.93	125 (33)	88 (33)	37 (33)	
Neutrophil-to-monocyte ratio, *n* (%)				0.942
<8.3	126 (33)	88 (33)	38 (34)	
8.3–10.0	125 (33)	87 (33)	38 (34)	
>10.0	126 (33)	90 (34)	36 (32)	
Neutrophil-to-HDL ratio, *n* (%)				0.411
<3.32	126 (33)	83 (31)	43 (38)	
3.32–4.90	126 (33)	91 (34)	35 (31)	
>4.90	125 (33)	91 (34)	34 (30)	
Lymphocyte-to-HDL ratio, *n* (%)				0.551
<1.10	126 (33)	91 (34)	35 (31)	
1.10–1.61	126 (33)	84 (32)	42 (38)	
>1.61	125 (33)	90 (34)	35 (31)	
Monocyte-to-HDL ratio, *n* (%)				0.551
<0.34	126 (33)	91 (34)	35 (31)	
0.34–0.51	126 (33)	84 (32)	42 (38)	
>0.51	125 (33)	90 (34)	35 (31)	
IL-6 (pg/mL), *n* (%)				0.533
<6.69	126 (33)	85 (32)	41 (37)	
6.69–12.99	126 (33)	93 (35)	33 (29)	
>12.99	125 (33)	87 (33)	38 (34)	
TNF-α (pg/mL), *n* (%)				0.294
<9.17	126 (33)	89 (34)	37 (33)	
9.17–16.90	126 (33)	94 (35)	32 (29)	
>16.90	125 (33)	82 (31)	43 (38)	
VCAM-1 (ng/mL), *n* (%)				0.302
<613.22	126 (33)	95 (36)	31 (28)	
613.22–873.64	126 (33)	86 (32)	40 (36)	

(Continued)

TABLE 1 (Continued)

Variables	Total (n = 377)	training set (n = 265)	validation set (n = 112)	p-value
>873.64	125 (33)	84 (32)	41 (37)	
LP-PLA2 (ng/mL), n (%)				0.79
<203.60	126 (33)	90 (34)	36 (32)	
203.60–307.46	126 (33)	90 (34)	36 (32)	
>307.46	125 (33)	85 (32)	40 (36)	
CD40L (ng/mL), n (%)				0.536
<2.01	126 (33)	84 (32)	42 (38)	
2.01–4.54	126 (33)	90 (34)	36 (32)	
>4.54	125 (33)	91 (34)	34 (30)	
E-selectin (ng/mL), n (%)				0.409
<23.75	126 (33)	87 (33)	39 (35)	
23.75–48.32	126 (33)	94 (35)	32 (29)	
>48.32	125 (33)	84 (32)	41 (37)	
ADMA (μmol/L), n (%)				0.334
<0.52	126 (33)	83 (31)	43 (38)	
0.52–0.73	126 (33)	89 (34)	37 (33)	
>0.73	125 (33)	93 (35)	32 (29)	
vWF (%), n (%)				0.562
<97.47	126 (33)	85 (32)	41 (37)	
97.47–149.19	126 (33)	88 (33)	38 (34)	
>149.19	125 (33)	92 (35)	33 (29)	
ICAM-1 (ng/mL), n (%)				0.605
<182.04	130 (34)	88 (33)	42 (38)	
182.04–199.66	27 (7)	18 (7)	9 (8)	
>199.66	220 (58)	159 (60)	61 (54)	

This table presents a comprehensive comparison of potential predictive factors for cognitive impairment in a cohort of 377 patients with cerebral small vessel disease (CSVD). Divided into a training set comprising 265 patients for model development and a validation set consisting of 112 patients for efficacy assessment, the data encompass a range of clinical characteristics and inflammatory biomarkers. These include gender, smoking and drinking habits, hypertension, diabetes, hyperlipidemia, along with inflammatory markers like white blood cell count, platelet count, neutrophil count, and specific indicators closely related to CSVD such as IL-6, TNFα, and VCAM-1. The objective of this table is to demonstrate the comparability of the training and validation sets in these potential predictors, ensuring the generalizability and reliability of the predictive model's outcomes. Statistical differences between the sets are indicated by p-values, with $p < 0.05$ denoting significant disparities.

3.4 Predictive model development

In this study, a nomogram was developed to predict the probability of cognitive impairment in patients with CSVD, based on four identified predictive factors: Hypertension, CSVD Burden, ApoA1 levels, and Age (Figure 3). Each factor contributes to an individualized risk score, calculated using the nomogram, which correlates with the likelihood of cognitive decline. This tool provides clinicians with a concise and quantifiable method to assess risk and tailor patient management strategies effectively.

3.5 Evaluation of the predictive nomogram

The predictive accuracy of our CSVD cognitive impairment nomogram was assessed using the area under the receiver operating characteristic (AUC-ROC) curve, with the training set demonstrating an AUC of 0.866 (95% CI: 0.823–0.909), and the validation set showing an AUC of 0.852 (95% CI: 0.781–0.923). These results indicate excellent discriminative ability across both datasets (Figures 4A,B). Calibration curves closely align with the ideal line in both training (Figure 5A) and validation (Figure 5B) sets, suggesting that the nomogram's predicted probabilities of cognitive impairment are accurate. Decision curve analysis (DCA) for both datasets confirms the model's clinical usefulness, with the net benefit substantially outweighing the treat-all or treat-none strategies (Figures 6A,B). In application, the nomogram showed high sensitivity and specificity, further evidencing its robustness. In the training set, the model exhibited a sensitivity of 75.3% and a specificity of 79.7% (Figure 4A). In contrast, in the validation set, the model showed a sensitivity of 76.9% and a specificity of 74.0% (Figure 4B). For the training set, the nomogram achieved a PPV of 87.32% and an NPV of 63.47%, and for the validation set, the PPV was 84.60% and NPV was 63.30%. These additional metrics further reinforce the nomogram's strong and consistent performance, advocating its potential utility in clinical practice for risk stratification and management of CSVD-associated cognitive impairment.

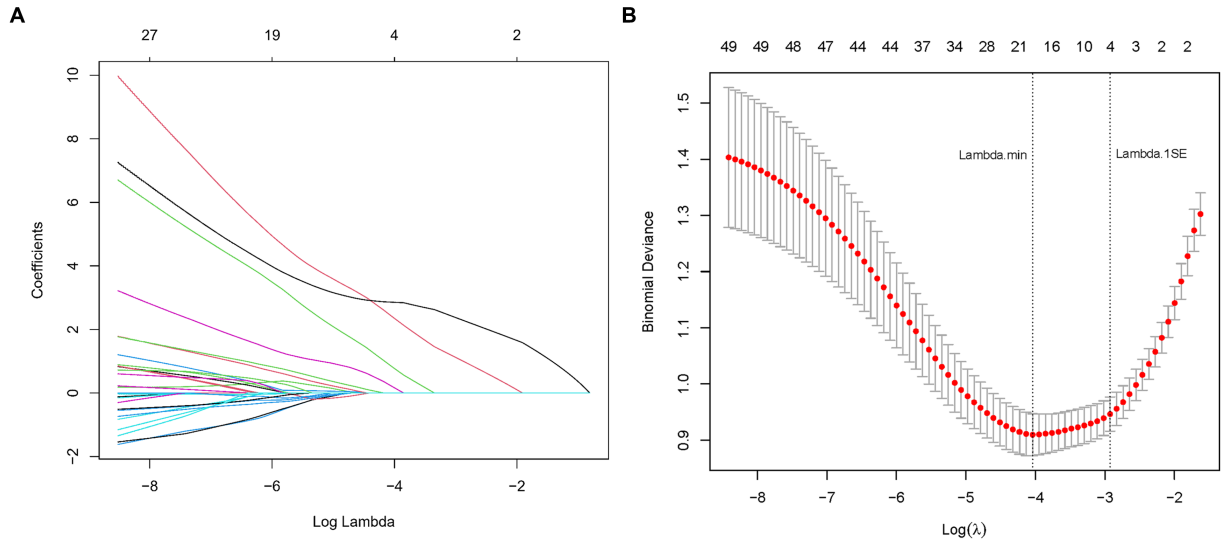


FIGURE 2
(A) LASSO coefficient profiles for CSVD cognitive impairment predictors. This figure shows the coefficient profiles of the predictors in the LASSO regression model. The *x*-axis represents the log of the regularization parameter (lambda), and the *y*-axis represents the coefficient values of the predictors. Each colored line represents a different predictor. As lambda increases (moving from left to right), the coefficients shrink towards zero, indicating the regularization effect of LASSO. The numbers at the top indicate the number of predictors included in the model for each lambda value. **(B)** Optimal lambda selection in LASSO model for CSVD cognitive impairment predictors. The graph demonstrates the selection of the optimal lambda value using cross-validation. The *x*-axis represents the log of lambda, and the *y*-axis represents the binomial deviance (a measure of model error). The red dots represent the mean binomial deviance for each lambda value, with error bars showing the standard error. The vertical dashed lines indicate the lambda values chosen by cross-validation: the left line represents the lambda that minimizes the binomial deviance, while the right line represents the largest lambda within one standard error of the minimum deviance, providing a more regularized and simpler model.

TABLE 2 **Coefficients and lambda.1SE value of the LASSO regression.**

Variable	Coefficients	Lambda.1SE
Hypertension	0.311	0.028
CSVD.Burden	0.0314	
ApoA1	−0.008	
Age	0.3131	

This table lists the coefficients and lambda.1SE values for key variables including hypertension, CSVD burden, ApoA1, and age in the LASSO regression model for cognitive impairment prediction in CSVD.

TABLE 3 **Binary logistic regression analysis.**

	B	SE	OR	CI	*Z*	*p*-value
Hypertension	3.21	0.557	1.78	1.32–3.84	3.767	<0.001
CSVD.Burden	0.602	0.203	1.83	1.23–2.72	2.97	0.003
ApoA1	−0.443	0.209	0.64	0.43–0.97	−2.126	0.034
Age	3.758	0.656	1.85	1.55–3.55	3.732	<0.001

This table shows key predictors for cognitive impairment in CSVD, including regression coefficients, odds ratios, and their statistical significance.

4 Discussion

The development of our predictive model for cognitive impairment in CSVD is a significant stride in neurovascular research, bridging a gap in early diagnostic methodologies. By integrating a diverse range of variables-demographic, clinical, neuroimaging, and biomarkers-our study sheds new light on the intricate relationship between CSVD and cognitive decline. This approach not only enhances our understanding of CSVD's impact on cognitive functions but also introduces a valuable tool for early detection and intervention, marking a crucial step forward in addressing this complex condition.

Our study contributes to the evolving landscape of CSVD research by presenting a comprehensive predictive model for cognitive impairment. This model integrates a wide array of variables, akin to the multifaceted approaches seen in recent literature. The deep learning approach of Duan et al. (14) in segmenting CSVD features on imaging contrasts with our model that combines clinical, biochemical, and imaging data for a more rounded prediction. Egle et al. (15) emphasized diffusion tensor imaging's role in predicting dementia, a perspective that complements our model's inclusion of imaging alongside other clinical variables. Our work resonates with the findings of Jiménez-Balado et al. (16) regarding the predictive value of blood pressure monitoring in CSVD. However, our approach is more comprehensive, encompassing a broader range of indicators. Similarly, while Li et al. (17) explored machine learning models for dementia prediction in CSVD, our study adds to this by leveraging both advanced statistical and machine learning techniques, emphasizing the integration of diverse data types. Furthermore, Liu et al. (9) highlighted the combination of Aβ42 levels and total CSVD scores in predicting cognitive impairment. Our model expands upon this by integrating these biomarkers into a broader predictive framework. Lastly, van Uden et al. (18) and Zhang et al. (10) focused on specific imaging and clinical parameters. Our research builds upon these studies by offering a more comprehensive model that incorporates their insights into a broader predictive framework. In summary, our study not only aligns with current research trends in CSVD but also extends them by providing a more holistic approach

10.3389/fneur.2024.1373306

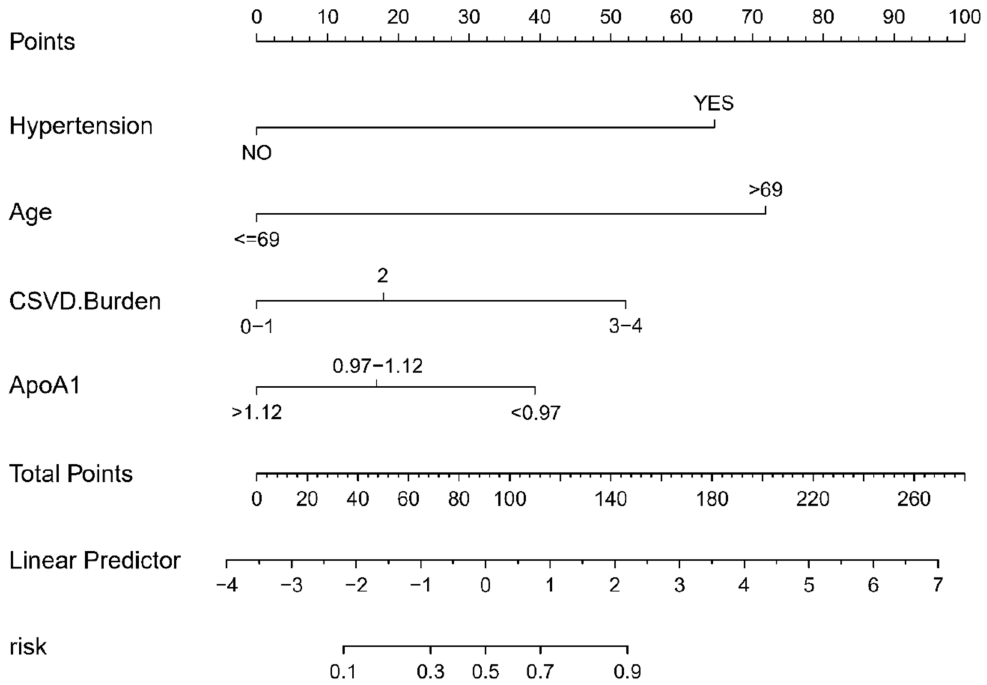


FIGURE 3
Nomogram for CSVD cognitive impairment risk prediction. This nomogram, based on logistic regression analysis, visually represents the contribution of each predictor to the overall risk of cognitive impairment in CSVD. Each predictor is assigned a score, and the total score corresponds to a probability of cognitive impairment. The predictors included are hypertension, CSVD burden, ApoA1 levels, and age.

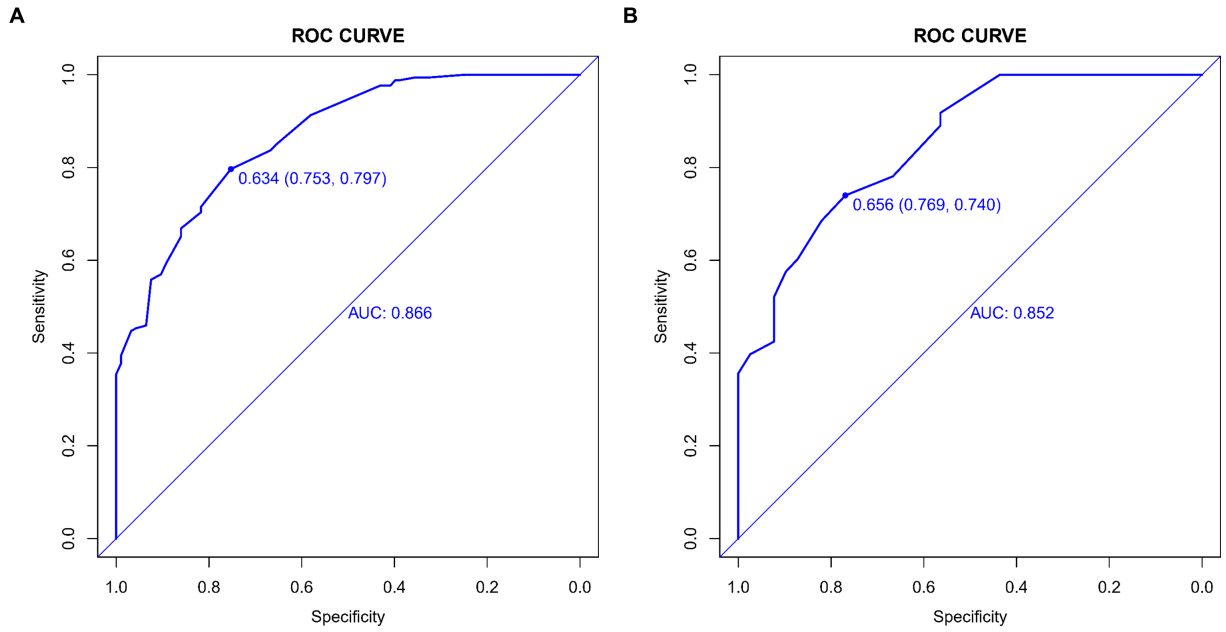


FIGURE 4
(A,B) ROC curves for CSVD cognitive impairment prediction model. These curves display the sensitivity and specificity of the model in both the training and validation sets, illustrating its diagnostic accuracy. The *x*-axis represents specificity, and the y-axis represents sensitivity. The area under the curve (AUC) values are shown for both sets, indicating the model's performance. **(A)** ROC curve for the training set, showing the model's performance with an AUC value of 0.866. **(B)** ROC curve for the validation set, showing the model's performance with an AUC value of 0.852.

to the prediction of cognitive impairment. This comprehensive model could significantly enhance clinical decision-making and patient management in CSVD.

Age is acknowledged as a critical predictor for the progression of CSVD and associated cognitive impairments. The study by Hamilton et al. (19) corroborates this, demonstrating a significant correlation

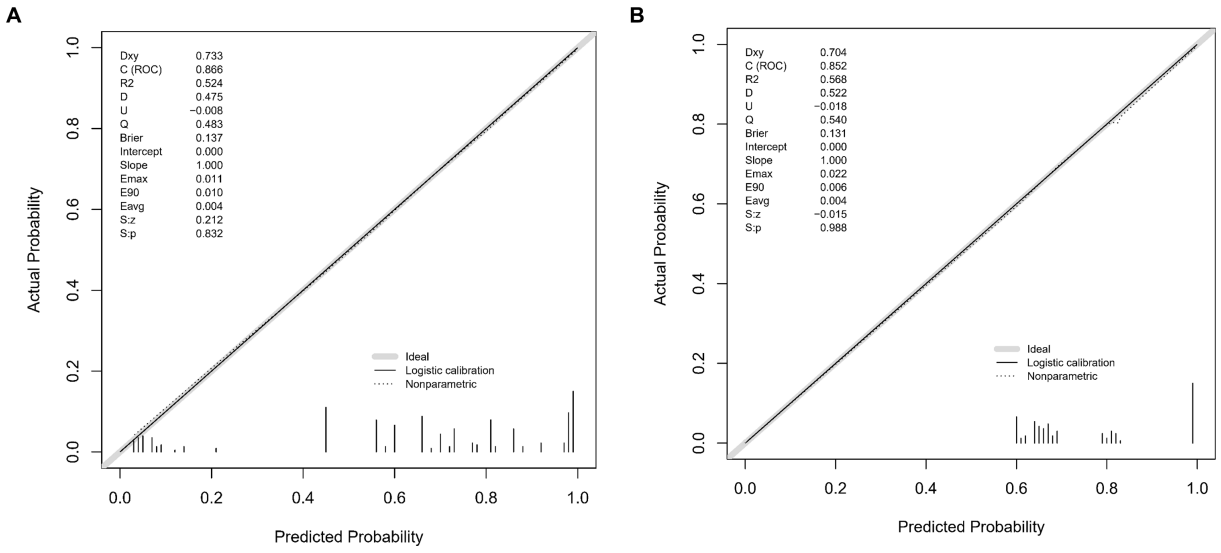


FIGURE 5
(A,B) Calibration plots for CSVD cognitive impairment prediction. These plots compare the predicted probabilities of cognitive impairment with the actual outcomes, demonstrating the model's calibration accuracy in both the training and validation sets. The *x*-axis represents the predicted probability, and the *y*-axis represents the actual probability. The diagonal line represents perfect calibration, where predicted probabilities exactly match the actual outcomes. **(A)** Calibration plot for the training set. The plot shows how well the predicted probabilities agree with the actual probabilities in the training data. The grey line represents the ideal calibration, the solid black line shows the logistic calibration, and the dotted line represents the nonparametric calibration. **(B)** Calibration plot for the validation set. Similar to **A**, this plot shows the agreement between predicted and actual probabilities in the validation data. The grey line represents the ideal calibration, the solid black line shows the logistic calibration, and the dotted line represents the nonparametric calibration.

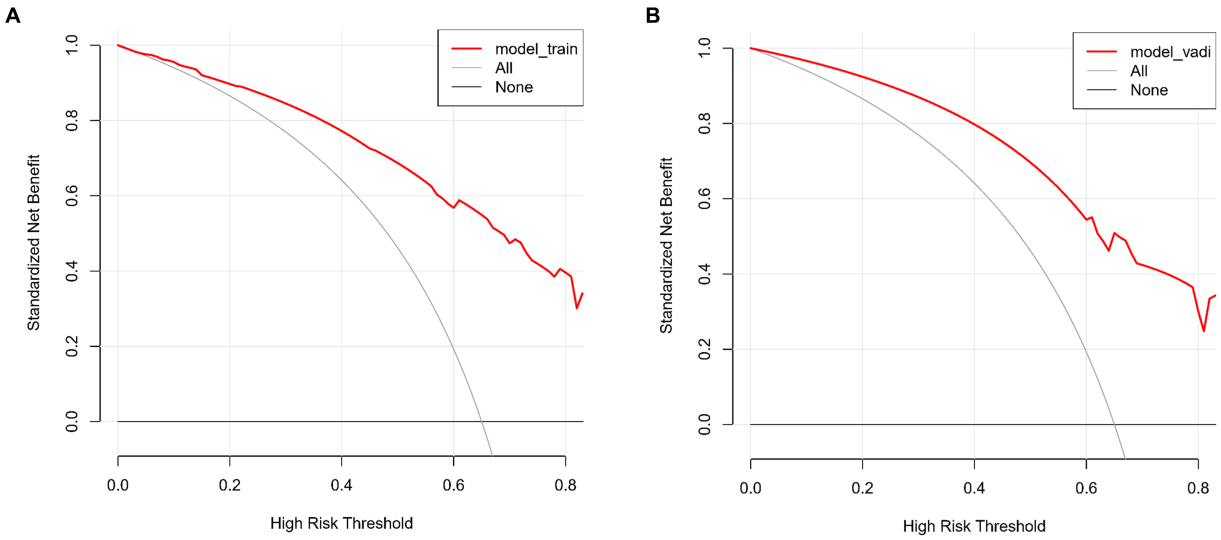


FIGURE 6
(A,B) Decision curve analysis for CSVD cognitive impairment model. These curves assess the clinical usefulness of the model by comparing the net benefits of different treatment strategies in both the training and validation sets. The *x*-axis represents the high-risk threshold, and the *y*-axis represents the standardized net benefit. **(A)** Decision curve for the training set. The red line represents the net benefit of the predictive model for the training data. The grey line, which slopes downward, represents the net benefit assuming all patients are treated. The black line, which is horizontal, represents the net benefit assuming no patients are treated. **(B)** Decision curve for the validation set. Similar to **A**, the red line represents the net benefit of the predictive model for the validation data. The grey line, which slopes downward, represents the net benefit assuming all patients are treated. The black line, which is horizontal, represents the net benefit assuming no patients are treated.

between the total burden of CSVD and a decline in cognitive abilities in the elderly. This finding not only validates the importance of age as a predictive factor but also underscores the heightened risk for cognitive health deterioration with advancing age. Hypertension is a primary contributor to the development of CSVD and subsequent cognitive decline. The research conducted by Amier et al. (20) revealed

a substantial association between markers of hypertensive exposure, as evident in cardiovascular MRI, and both CSVD and cognitive impairments. Additionally, the study by Hainsworth et al. (21) accentuates the close link between hypertension and small vessel disease, aligning well with our findings and underscoring the necessity to consider hypertension's role in clinical predictive models. The imaging burden of CSVD is significantly associated with declines in cognitive function. Hosoya et al. (22) observed an independent association between imaging markers of CSVD and reductions in global cognitive function and attention. Our research echoes this observation, highlighting the crucial role of CSVD imaging burden in predicting cognitive impairments. Our study further identifies ApoA1 as a novel negative predictor for cognitive impairment in CSVD, aligning with emerging research in neurodegenerative conditions. Studies by Choi et al. (23) and Slot et al. (24) reported similar associations between ApoA1 levels and cognitive decline, suggesting ApoA1's potential role in amyloid-independent neurodegeneration. Moreover, research by Das et al. (25) and Deng et al. (26) underscore ApoA1's neuroprotective effects in Parkinson's disease, while Rao et al. (27) highlight the complex interaction between ApoA1 levels and genetic factors in cognitive impairment. These findings collectively corroborate our results, emphasizing ApoA1's significance in diagnosing and treating CSVD-related cognitive impairment.

In our study, we have elucidated several pivotal findings that hold significant implications for understanding the development of cognitive impairment in the context of CSVD. Firstly, we have reaffirmed hypertension and advanced age as the principal risk factors for cognitive impairment in CSVD. Both factors exhibit a substantial association with an increased risk of cognitive impairment, underscoring the imperative need for clinicians to closely monitor cognitive functions in patients exhibiting these characteristics. Additionally, our model underscores the critical importance of the cumulative burden of CSVD-related lesions (CSVD burden) and serum levels of apolipoprotein A1 (ApoA1) in predicting cognitive function. These findings provide valuable clinical cues to assist physicians in identifying high-risk individuals and devising tailored management strategies.

The results of our investigation hold profound clinical utility. The predictive model we have developed, based on CSVD risk factors, serves as a valuable tool for the early identification of individuals at risk of developing cognitive impairment. This not only facilitates timely intervention and treatment but also has the potential to enhance the quality of life for affected individuals. Our model offers a personalized risk assessment tool that guides clinicians in formulating precise treatment plans, while also empowering patients with information about their individual risks. Despite the significant achievements of our study, several avenues for future research merit exploration. Firstly, we advocate for further research into the biomarkers of CSVD-related cognitive impairment to refine risk prediction. Furthermore, as neuroimaging and molecular biology techniques continue to advance, we encourage the integration of these advanced technologies into a more comprehensive predictive model for CSVD. Additionally, longitudinal studies with extended follow-up periods are warranted to gain deeper insights into the progression and trajectory of cognitive impairment in CSVD.

Our study is not without limitations. Firstly, its cross-sectional design precludes the establishment of causal relationships. Although our model demonstrates excellent predictive performance for CSVD-related cognitive impairment, further validation in larger multicenter cohorts is warranted. Additionally, the potential for selection bias may exist, as all participants were sourced from a single hospital. Despite the inclusion of a comprehensive array of clinical, biochemical, and imaging data, there may be other unexplored factors that influence the risk of cognitive impairment in CSVD. Moreover, the sample size of this study is relatively small. Future research should include a larger sample size to enhance the reliability and generalizability of the findings.

In summary, our study presents a comprehensive predictive model that holds significant promise in forecasting the risk of cognitive impairment in the context of CSVD. This model integrates clinical, biochemical, and imaging data, furnishing clinicians with a powerful tool to identify high-risk individuals and devise personalized management strategies. While limitations and avenues for future research exist, we believe that this study marks a pivotal advancement in the early diagnosis and intervention of CSVD-related cognitive impairment.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Affiliated Hospital of Hebei University. The committee is affiliated with Hebei University, located in Baoding, China. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. YG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. L-tL: Conceptualization, Formal analysis, Methodology, Software, Supervision, Writing – review & editing. Y-dH: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing. LL: Investigation, Methodology, Validation, Writing – review & editing. NJ: Data curation, Software, Writing – review & editing. Y-jC: Formal analysis, Software, Writing – review & editing. Y-nM: Investigation, Methodology, Writing – review & editing. YJ: Conceptualization, Formal analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by the Hospital fund of Affiliated Hospital of Hebei University, Grant/Award Number: 2022QC43 and Hebei Province Medical Science Research Project, Grant/Award Number: 20230200.

Acknowledgments

We extend our heartfelt gratitude to neuroimaging experts Yan Hou and Huan Zhou for their invaluable expertise and contributions to this study. Their insights and assistance were instrumental in the advancement of our research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Azarpazhooh MR, Avan A, Cipriano LE, Munoz DG, Sposato LA, Hachinski V. Concomitant vascular and neurodegenerative pathologies double the risk of dementia. *Alzheimers Dement*. (2018) 14:148–56. doi: 10.1016/j.jalz.2017.07.755

2. Debette S, Schilling S, Duperron M-G, Larsson SC, Markus HS. Clinical significance of magnetic resonance imaging markers of vascular brain injury: a systematic review and meta-analysis. *JAMA Neurol*. (2019) 76:81–94. doi: 10.1001/jamaneurol.2018.3122

3. Pan Y, Li H, Wardlaw JM, Wang Y. A new dawn of preventing dementia by preventing cerebrovascular diseases. *BMJ*. (2020) 371:m3692. doi: 10.1136/bmj.m3692

4. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol*. (2010) 9:689–701. doi: 10.1016/S1474-4422(10)70104-6

5. Wardlaw JM, Smith C, Dichgans M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurol*. (2013) 12:483–97. doi: 10.1016/S1474-4422(13)70060-7

6. Wardlaw JM, Smith C, Dichgans M. Small vessel disease: mechanisms and clinical implications. *Lancet Neurol*. (2019) 18:684–96. doi: 10.1016/S1474-4422(19)30079-1

7. Kapasi A, DeCarli C, Schneider JA. Impact of multiple pathologies on the threshold for clinically overt dementia. *Acta Neuropathol*. (2017) 134:171–86. doi: 10.1007/s00401-017-1717-7

8. Das AS, Regenhardt RW, Vernooij MW, Blacker D, Charidimou A, Viswanathan A. Asymptomatic cerebral small vessel disease: insights from population-based studies. *J Stroke*. (2019) 21:121–38. doi: 10.5853/jos.2018.03608

9. Liu J, Zhao W, Gui Q, Zhang Y, Guo Z, Liu W. Addition of Aβ42 to total cerebral small vessel disease score improves the prediction for cognitive impairment in cerebral small vessel disease patients. *Neuropsychiatr Dis Treat*. (2021) 17:195–201. doi: 10.2147/NDT.S289357

10. Zhang L, Gao F, Zhang Y, Hu P, Yao Y, Zhang Q, et al. Analysis of risk factors for the development of cognitive dysfunction in patients with cerebral small vessel disease and the construction of a predictive model. *Front Neurol*. (2022) 13:944205. doi: 10.3389/fneur.2022.944205

11. Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol*. (2013) 12:822–38. doi: 10.1016/S1474-4422(13)70124-8

12. Staals J, Makin SDJ, Doubal FN, Dennis MS, Wardlaw JM. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology*. (2014) 83:1228–34. doi: 10.1212/WNL.0000000000000837

13. Duering M, Biessels GJ, Brodtmann A, Chen C, Cordonnier C, de Leeuw FE, et al. Neuroimaging standards for research into small vessel disease-advances since 2013. *Lancet Neurol*. (2023) 22:602–18. doi: 10.1016/S1474-4422(23)00131-X

14. Duan YY, Shan W, Liu LY, Wang Q, Wu Z, Liu P, et al. Primary categorizing and masking cerebral small vessel disease based on "deep learning system". *Front Neuroinform*. (2020) 14:17. doi: 10.3389/fninf.2020.00017

15. Egle M, Hilal S, Tuladhar AM, Pirpamer L, Hofer E, Duering M, et al. Prediction of dementia using diffusion tensor MRI measures: the OPTIMAL collaboration. *J Neurol Neurosurg Psychiatry*. (2022) 93:14–23. doi: 10.1136/jnnp-2021-326571

16. Jiménez-Balado J, Riba-Llena I, Maisterra O, Pizarro J, Palasí A, Pujadas F, et al. Ambulatory blood pressure levels in the prediction of progression of cerebral small vessel disease. *J Am Geriatr Soc*. (2020) 68:2232–9. doi: 10.1111/jgs.16568

17. Li R, Harshfield EL, Bell S, Burkhart M, Tuladhar AM, Hilal S, et al. Predicting incident dementia in cerebral small vessel disease: comparison of machine learning and traditional statistical models. *Cereb Circ Cogn Behav*. (2023) 5:100179. doi: 10.1016/j.cccb.2023.100179

18. van Uden IWM, van der Holst HM, Tuladhar AM, van Norden A, de Laat KF, Rutten-Jacobs LC, et al. White matter and hippocampal volume predict the risk of dementia in patients with cerebral small vessel disease: the RUN DMC study. *J Alzheimers Dis*. (2016) 49:863–73. doi: 10.3233/JAD-150573

19. Hamilton OKL, Cox SR, Okely JA, Conte F, Ballerini L, Bastin ME, et al. Cerebral small vessel disease burden and longitudinal cognitive decline from age 73 to 82: the Lothian Birth Cohort 1936. *Transl Psychiatry*. (2021) 11:376. doi: 10.1038/s41398-021-01495-4

20. Amier RP, Marcks N, Hooghiemstra AM, Nijveldt R, van Buchem M, de Roos A, et al. Hypertensive exposure markers by MRI in relation to cerebral small vessel disease and cognitive impairment. *J Am Coll Cardiol Img*. (2021) 14:176–85. doi: 10.1016/j.jcmg.2020.06.040

21. Hainsworth AH, Markus HS, Schneider JA. Cerebral small vessel disease, hypertension, and vascular contributions to cognitive impairment and dementia. *Hypertension*. (2024) 81:75–86. doi: 10.1161/HYPERTENSIONAHA.123.19943

22. Hosoya M, Toi S, Seki M, Saito M, Hoshino T, Yoshizawa H, et al. Association between total cerebral small vessel disease score and cognitive function in patients with vascular risk factors. *Hypertens Res*. (2023) 46:1326–34. doi: 10.1038/s41440-023-01244-8

23. Choi HJ, Seo EH, Yi D, Sohn BK, Choe YM, Byun MS, et al. Amyloid-independent amnestic mild cognitive impairment and serum apolipoprotein A1 levels. *Am J Geriatr Psychiatry*. (2016) 24:144–53. doi: 10.1016/j.jagp.2015.06.004

24. Slot RER, van Harten AC, Kester MI, Jongbloed W, Bouwman FH, Teunissen CE, et al. Apolipoprotein A1 in cerebrospinal fluid and plasma and progression to Alzheimer's disease in non-demented elderly. *J Alzheimers Dis*. (2017) 56:687–97. doi: 10.3233/JAD-151068

25. Das S, Charan R, Skrinak RT, Zhu, Y, Lee, M, Kumar, S., et al. ApolipoproteinA1 has a protective role in the pathogenesis of Parkinson's disease. *Ann Neurol*. (2022) 92:S156–7. doi: 10.1002/ana.26484

26. Deng X, Saffari SE, Ng SYE, Chia N, Tan JY, Choi X, et al. Blood lipid biomarkers in early Parkinson's disease and Parkinson's disease with mild cognitive impairment. *J Parkinsons Dis*. (2022) 12:1937–43. doi: 10.3233/JPD-213135

27. Rao W, Zhang Y, Li K, Zhang XY. Association between cognitive impairment and apolipoprotein A1 or apolipoprotein B levels is regulated by apolipoprotein E variant rs429358 in patients with chronic schizophrenia. *Aging*. (2021) 13:16353–66. doi: 10.18632/aging.203161

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 26 June 2024
DOI 10.3389/fneur.2024.1386802

Check for updates

OPEN ACCESS

EDITED BY
Tarun Singh,
University of Michigan, United States

REVIEWED BY
Joshua Cohn,
Fox Chase Cancer Center, United States
Thomas Hsueh,
Taipei City Hospital, Taiwan

*CORRESPONDENCE
Yusuf Mehkri
✉ yusuf.mehkri@neurosurgery.ufl.edu

†These authors share first authorship

RECEIVED 16 February 2024
ACCEPTED 11 June 2024
PUBLISHED 26 June 2024

CITATION
Malnik SL, Porche K, Mehkri Y, Yue S, Maciel CB, Lucke-Wold BP, Robicsek SA, Decker M and Busl KM (2024) Leveraging machine learning to develop a postoperative predictive model for postoperative urinary retention following lumbar spine surgery. *Front. Neurol.* 15:1386802. doi: 10.3389/fneur.2024.1386802

COPYRIGHT
© 2024 Malnik, Porche, Mehkri, Yue, Maciel, Lucke-Wold, Robicsek, Decker and Busl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Leveraging machine learning to develop a postoperative predictive model for postoperative urinary retention following lumbar spine surgery

Samuel L. Malnik[1†], Ken Porche[2†], Yusuf Mehkri[2*], Sijia Yue[3], Carolina B. Maciel[4], Brandon P. Lucke-Wold[2], Steven A. Robicsek[5], Matthew Decker[2] and Katharina M. Busl[4]

[1]Department of Neurosurgery, Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix, AZ, United States, [2]Lillian S. Wells Department of Neurosurgery, University of Florida, Gainesville, FL, United States, [3]Department of Biostatistics, University of Florida, Gainesville, FL, United States, [4]Departments of Neurology and Neurosurgery, University of Florida, Gainesville, FL, United States, [5]Department of Anesthesiology, University of Florida, Gainesville, FL, United States

Introduction: Postoperative urinary retention (POUR) is the inability to urinate after a surgical procedure despite having a full bladder. It is a common complication following lumbar spine surgery which has been extensively linked to increased patient morbidity and hospital costs. This study hopes to development and validate a predictive model for POUR following lumbar spine surgery using patient demographics, surgical and anesthesia variables.

Methods: This is a retrospective observational cohort study of 903 patients who underwent lumbar spine surgery over the period of June 2017 to June 2019 in a tertiary academic medical center. Four hundred and nineteen variables were collected including patient demographics, ICD-10 codes, and intraoperative factors. Least absolute shrinkage and selection operation (LASSO) regression and logistic regression models were compared. A decision tree model was fitted to the optimal model to classify each patient's risk of developing POUR as high, intermediate, or low risk. Predictive performance of POUR was assessed by area under the receiver operating characteristic curve (AUC-ROC).

Results: 903 patients were included with average age 60 ± 15 years, body mass index of 30.5 ± 6.4 kg/m^2, 476 (53%) male, 785 (87%) white, 446 (49%) involving fusions, with average 2.1 ± 2.0 levels. The incidence of POUR was 235 (26%) with 63 (7%) requiring indwelling catheter placement. A decision tree was constructed with an accuracy of 87.8%.

Conclusion: We present a highly accurate and easy to implement decision tree model which predicts POUR following lumbar spine surgery using preoperative and intraoperative variables.

KEYWORDS

lumbar surgery, machine learning, postoperative complications, risk factors, urinary catheterization, urinary retention

1 Introduction

Postoperative urinary retention (POUR) refers to a patient's inability to completely empty their distended bladder following surgery. POUR is a common complication across all surgical specialties with an incidence of 5%–70% (1, 2). Following spine surgery, average rates of POUR range from 5 to 38% depending on the definition of POUR, study population, and surgical characteristics (1, 3–8). The occurrence of POUR leads to discomfort and the potential need for catheterization, factors that overtly impact patient well-being. POUR has also been extensively linked to increased risk for serious complications such as urinary tract infection, sepsis, increased length of stay, higher medical costs, and increased rates of readmission to the hospital (4, 5, 9–11). In addition to immediate patient well-being and comfort, POUR was found to lower patient satisfaction, with patients who experienced POUR being less likely to be satisfied with spine surgery even at long-term follow up (11).

Several patient specific risk factors have been associated with the development of POUR following lumbar spine surgery with age and male sex being the most frequently described factors (4, 5, 9, 10, 12). Likewise, numerous surgical factors such as operative time, number of operative levels, and fusion/surgical instrumentation have been associated with POUR (4, 5, 11–13). While dozens of factors have been analyzed, few of these analyses have brought forth actionable plans for identifying patients at greatest risk for POUR outside of single variable analysis. These univariate approaches fail to adequately analyze the complex interactions of patient and surgical variables which limits their predictive accuracy.

Machine learning has become widely popularized in the spine surgery literature over the past decade with its application being put forward toward diagnosis of spinal conditions and prediction of surgical complications and outcomes (14). Previously, our group published a highly accurate model using preoperative variables to predict POUR through regression and neural network analysis (15); however, it did not account for intraoperative and perioperative variables during anesthesia, such as administration of narcotics, that have been demonstrated to affect a patient's likelihood to develop POUR (16–18). Herein, we present a machine learning comprehensive approach for identification and classification of patients at risk for POUR following lumbar spine surgery with patient, surgical and anesthesia variables. We hypothesize that the inclusion of a greater spectrum of variables will increase the fidelity of the predictive model. Practically, this would enable the surgical team to better identify patients at greatest risk for POUR, proactively adjust expectations, and arrange for proper monitoring and mitigating strategies.

2 Methods

2.1 Study design

We performed a retrospective review of consecutive patients who underwent spine surgery at our tertiary care academic medical center from June 2017 to June 2019. Patients were identified for inclusion in the database by query of CPT codes specific to lumbar spine operations: 22533, 22534, 22558, 22585, 22612, 22614, 22630, 22633, 22634, 63005, 63012, 63017, 63030, 63035, 63042, 63047, 63048, 63056, 63057. Patients were excluded if surgery was not done through the clinic setting, had surgery in a non-lumbar region (i.e., thoracic, or cervical level), or were <18 years old. Study design and data security methods were approved by our Institutional Review Board under protocol #201902403.

2.2 Identification of variables

The data were retrospectively collected from charted demographic information, nursing and anesthesia reports, and neurosurgical operative reports. Preoperative variables included age, body mass index (BMI), and pre-surgical use of opioids or urinary retention medication (i.e., 5-alpha reductase inhibitors and/or alpha inhibitors). International Classification of Diseases (ICD) codes preexisting the surgical visit were collected from electronic health record (EHR) as well as Epic's Care Everywhere® feature, a network connecting UF Health's EHR to hundreds of other EHRs utilizing the Epic system (Epic Systems Corporation, Verona, Wisconsin). Intraoperative and post-operative variables were chosen based on previous studies and clinical suspicion of relevance (1, 9, 11–13, 15, 17–22). Intraoperative surgical variables included duration of surgery, indwelling catheter use, type of surgery (discectomy, laminectomy, and/or fusion), type of fusion if relevant, pelvic screw placement, number of levels, use of minimally invasive techniques, and surgical approach. Intraoperative anesthesia variables included total intravenous fluid administration, total volume of blood products transfused, and all medications administered during the surgical procedure.

2.3 Definition of POUR

Patients were monitored in the neuroscience intensive care unit, post-anesthesia care unit, and neurosurgical floor unit for failure to void and distended or painful bladders. Indwelling urinary catheters were placed intraoperatively for cases with expected surgery duration exceeding 3 h. In the absence of indwelling catheters, urine volume was determined per standard of care with nurse-led bladder scanning. POUR was defined as the reinsertion of indwelling urinary catheter, or the need for straight catheterization for urine volumes exceeding 400 mL on bladder scan (23, 24). Bladder scan was done with ultrasound in standard fashion. Timing of postoperative removal of the indwelling urinary catheter occurred at the discretion of the surgeon.

3 Statistical analysis

3.1 Variable selection

Four hundred and nineteen variables were collected including patient characteristics, ICD-10 codes, and intraoperative factors. Only patients with complete data sets were included in the analysis. To set up a model for predicting POUR, variables were selected in two steps. In the first selection stage, all variables were subjected to univariate analysis to reveal patterns of association with POUR. Mann–Whitney U-tests were used for continuous and nominal variables while chi-square tests were used for categorical variables. Following this analysis, variables were selected depending on statistical significance and refined based on previous literature (2, 15). Then, a LASSO regression approach based on a penalized regression to obtain shrinkage estimators where only variables that did not shrink to 0 were kept.

The data were randomly split into training (80%) and validation sets (20%). The training set was used to develop models to predict POUR. The validation set was used to evaluate the performance of the prediction models that fitted from the training data.

3.2 Predictive modeling

In building the predictive models, a logistic regression model is first fitted to predict POUR using the selected variables. The area under the curve (AUC) on both training and validation dataset was assessed to show the performance. Then, the predicted probability of having POUR for all patients from training and validation set is calculated from the logistic regression model. Based on the distribution of outcomes found in prior modeling based on pre-operative risk factors, we defined the top 11% of the predicted probability as high risk, the 74% as intermediate risk and the last 15% as low risk (15). Using the risk levels as outcome, a decision tree model is fitted to classify each patient's risk level in the training set. Five-fold cross validation is utilized for hyper parameter tuning on minimum split and maximum depth. The accuracy of the decision tree is calculated from the validation set for performance evaluation. Brier score (measure of the accuracy of the probalistic prediction) was used to compare the forecasting ability of each aspect of the model, where the lower the score, the better the predictions are calibrated (25). All statistical analyses were performed using SAS statistical software.

4 Results

4.1 Clinical characteristics

Of the 1,387 patients enrolled via CPT codes, 362 were non-lumbar, 77 were found to be non-elective, and 45 patients had missing data as shown in Figure 1. Of 903 patients included in this study, the mean age was 59.5 ± 15.4 years, BMI of 30.5 ± 6.4 kg/m^2, 476 (53%) male, and 785 (87%) white. 24/903 (2.7%) had a history of UTI, and 27/903 (2.9%) had a history of retention. The incidence of POUR was 235 (26.1%) with 63 (7%) requiring indwelling urinary catheter placement. Patients who developed POUR were significantly older (62.2 ± 15.4 years vs. 58.5 ± 15.4 years, $p = 0.002$) but did not significantly differ with regards to BMI (30.6 ± 6.63 kg/m^2 vs. 30.3 ± 5.85 kg/m^2, $p = 0.488$), male sex (44.9% vs. 48.0%, $p = 0.414$), or white race (86.1% vs. 87.3%, $p = 0.659$). Differences in the rates of POUR based on preoperative clinical characteristics are shown in Figure 2. Patients who developed POUR were statistically more likely to have taken tamsulosin (+16.6%, $p = 0.050$) or opioids prior to surgery (+11.7%, $p < 0.002$), had an American Society of Anesthesiologist Physical Status Classification System (ASA) score > 2 (+11.2%, $p = 0.001$), and had a Charlson Comorbidity Index (CCI) > 1 (+10%, $p = 0.001$).

4.2 Surgical characteristics

The differences in rates of POUR based on surgical variables are shown in Figure 3. There were multiple significant surgical predictors of POUR. Rates of POUR were significantly higher in patients with surgeries involving fusion (+18.4%, $p < 0.001$) or laminectomy (+13.2%, $p < 0.001$). The rates of POUR in patients who underwent multilevel laminectomy (+22.1%, $p < 0.001$) and multilevel fusion (+24.1%, $p < 0.001$) were higher. Intraoperative indwelling urinary catheter placement (+20.1%, $p < 0.001$) was a strong predictor of POUR. Similarly, there was a significant difference in the likelihood to develop POUR in patients who underwent surgery involving posterolateral fusion (+18.8%, $p < 0.001$), pelvic screw placement (+15.9%, $p = 0.014$) or interbody fusion (+9%, $p < 0.003$). Conversely, rates of POUR were significantly lower in patients whose surgery

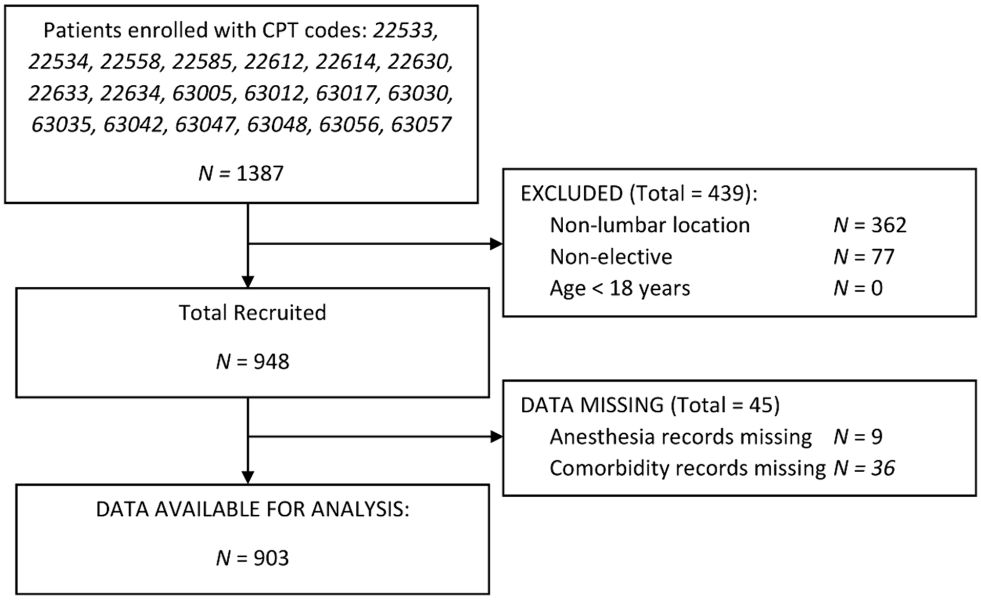


FIGURE 1
Flow diagram. CPT, Current procedural terminology.

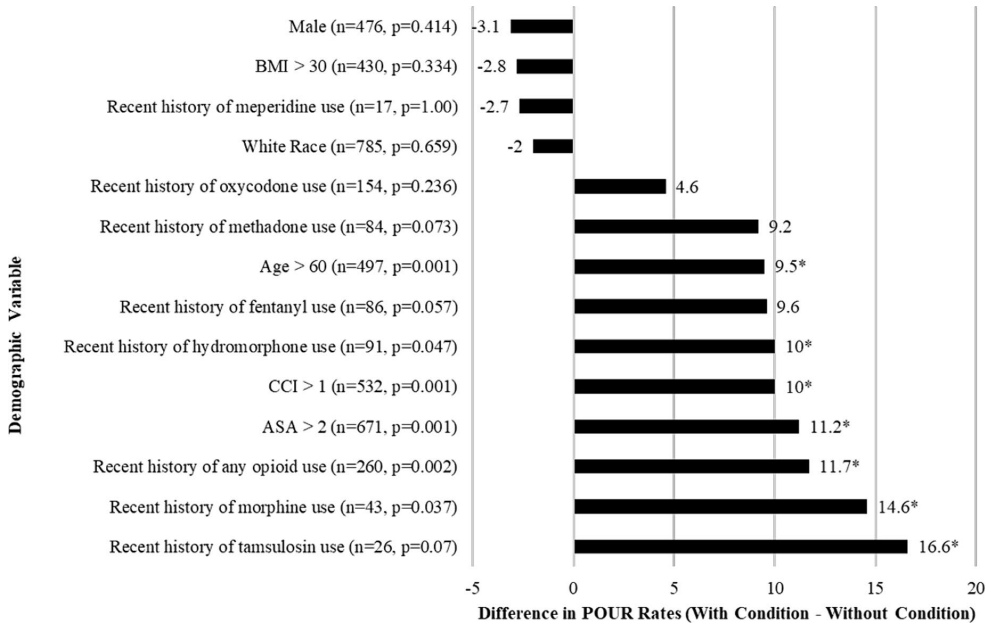


FIGURE 2
Bar graph of the differences in the rates of POUR based on preoperative clinical characteristics for patients who underwent lumbar spine surgery. Frequency (n) and *p*-values comparing those who did and did not develop POUR. Asterisk (*) indicates $p < 0.05$ in chi-square tests. BMI, Body mass index.

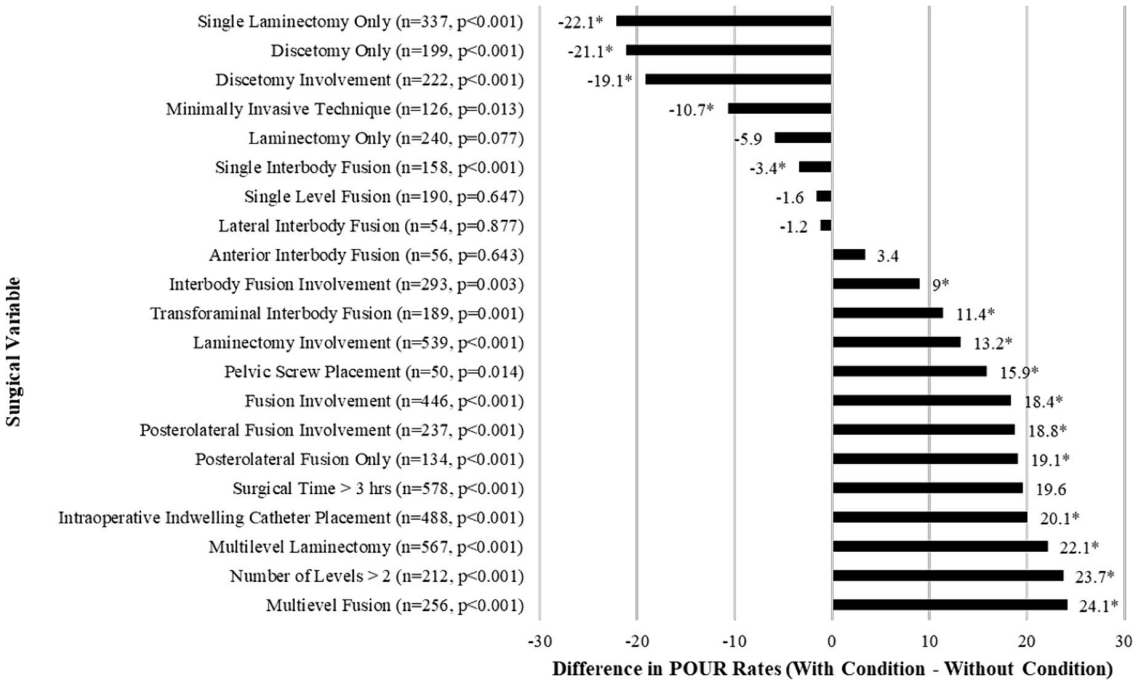


FIGURE 3
Bar graph of the differences in the rates of POUR based on categorical surgical variables for patients who underwent lumbar spine surgery. Frequency (n) and *p*-values comparing those who did and did not develop POUR. Asterisk (*) indicates $p < 0.05$ in chi-square test.

included discectomy only (−22.1%, $p < 0.001$) or involved discectomy (−19.1%, $p < 0.001$). Similarly, rates of POUR were significantly lower in patients who underwent minimally invasive technique operations (−10.7%, $p < 0.001$).

4.3 Anesthesia characteristics

A total of 69 variables were extracted and analyzed from intraoperative charts including muscle relaxants, reversal agents,

vasopressors, antihypertensives, antibiotics, neuromuscular agents, sedatives, analgesics (opioids and non-opioid), intravenous fluids, and blood product transfusions. The average amount of 25 anesthesia variables were found to be significantly different between the groups of patients (Table 1). Patients who developed POUR had a significantly longer average surgical time (310 ± 147 min vs. 236 ± 130 min, $p < 0.001$), received greater volume of intravenous fluids (3,000 ± 2,330 mL vs. 1,960 ± 1,520 mL, $p < 0.001$), and received greater oral morphine equivalents (OME) of intravenous opioids (21.3 ± 35.0 mg OME vs. 13.1 ± 28.4 mg OME, $p < 0.001$).

Following initial univariate analysis of patient, surgical and anesthesia-related factors, 94 variables were selected for LASSO regression of which 13 variables did not shrink to 0. The LASSO regression model achieved an AUC of 0.676 on the testing set on the receiver operating characteristic (ROC) curve (training set AUC 0.743). The AUC on the precision recall curve (PRC) were 0.332 and 0.560 for the testing and training sets, respectively. After the model selection step, 14 variables including patient, surgical and anesthesia factors were isolated and included in logistic regression (Table 2). The logistic regression outperformed the LASSO regression model with an AUC-ROC of 0.737 (training set AUC 0.768; Figure 4). The AUC-PRC for this model on the testing and training sets were 0.614 and 0.402, respectively. After hyper-parametric tuning of selected predictors from the LASSO regression model, a decision tree model was constructed (Figure 5). The accuracy for the final decision tree model was confirmed to be 87.8% on a 3-class confusion matrix (which reduces to 70.9% on a confusion matrix excluding the intermediate category), with sensitivity 91.3%, specificity 55.2%, positive predictive value 61.0%, and negative predictive value 89.2%. Brier score was noted to be 0.19.

5 Discussion

POUR is an incompletely understood but frequently encountered barrier to patient recovery and satisfaction following lumbar spine surgery occurring in 25% of patients. Its pathogenesis is thought to be related to several factors including anesthetic agents, perioperative medications, and postoperative pain, all of which can alter the complex urinary signaling pathway. Anesthetics can act centrally at the pontine micturition center and peripherally as smooth muscle relaxants to decrease bladder contractility (26). Surgical pain or inadequate pain control further stimulates the sympathetic nervous

TABLE 1 Selected anesthesia variables found to have statistically significant differences between the group of patients that developed POUR and the group of patients that did not develop POUR.

Variables—Mean (SD)*	POUR	No POUR	*p*-value	Correlation
Albumin (g)	17.7 (29.3)	7.5 (1.8)	<0.001	+
Calcium chloride	189 (711)	67.1 (342)	0.001	+
Calcium gluconate	189 (711)	71.5 (451)	0.014	+
Cefazolin	2,830 (2,060)	2,390 (1,720)	0.003	+
Dexamethasone	1.68 (3.43)	2.52 (3.98)	0.001	–
Ephedrine	12.2 (14.8)	9.18 (14.6)	0.005	+
Hydromorphone	0.397 (0.867)	0.262 (0.702)	0.031	+
Total IV Fluid Volume (mL)	3,000 (2,330)	1,960 (1,520)	<0.001	+
Ketorolac	1.91 (6.52)	4.49 (10.4)	<0.001	+
Methadone	3.96 (6.94)	2.41 (5.56)	0.001	+
Midazolam	0.536 (0.944)	0.805 (1.82)	0.003	–
Neostigmine	0.151 (0.784)	0.0419 (0.408)	0.043	+
Ondansetron	3.68 (1.27)	3.82 (1.17)	0.040	–
Oral Morphine Equivalents	21.3 (35.0)	13.1 (28.4)	0.001	+
Phenylephrine	6.39 (7.57)	2.99 (4.88)	<0.001	+
Plasma Transfusion (mL)	18.3 (134)	1.22 (23.8)	0.039	+
Plasmalyte (mL)	2,190 (1,480)	1,540 (1,210)	<0.001	+
Platelet Transfusion (mL)	6.94 (49.4)	0.753 (19.5)	0.038	+
Promethazine	0.0426 (0.460)	0.161 (1.15)	0.009	–
Propofol	948 (1,980)	645 (1,300)	0.019	+
RBC Transfusion (mL)	156 (546)	26.7 (167)	<0.001	+
Remifentanil	0.475 (1.40)	0.242 (1.35)	0.038	+
Rocuronium	87.0 (50.1)	75.9 (39.7)	0.006	+
Sufentanil	0.034 (0.068)	0.023 (0.045)	0.015	+
Surgery Time (min)	310 (147)	236 (130)	<0.001	+

Correlation indications relationship between variable and association with POUR. *Units in mg unless otherwise mentioned. SD, Standard Deviation; POUR, Postoperative urinary retention; IV, Intravenous; RBC, Red Blood Cell.

TABLE 2 Multivariate logistic regression analysis for the development of the POUR model.

Variable	Estimate	SE	Statistic	*p*-value
Age (years)	0.012	0.007	1.738	0.082
Arthrodesis—Z98.1	0.225	0.221	1.019	0.308
Cardiomegaly—I51.7	0.725	0.439	1.649	0.099
Constipation—K59.00	1.777	0.568	3.128	**0.002**
Discectomy Involvement	−0.389	0.317	−1.228	0.220
Ileus—K56.7	1.178	0.621	1.896	0.058
Intraoperative Foley	0.569	0.245	2.323	**0.020**
IV Fluid Volume (mL)	<0.001	<0.001	−0.579	0.562
Neostigmine (mg)	0.388	0.151	2.562	**0.010**
Number of Disc Levels	0.025	0.056	0.45	0.653
Phenylephrine (mg)	0.054	0.019	2.887	**0.004**
Pleural Effusion—J90	0.709	0.557	1.274	0.203
RBC Transfusion (mL)	0.001	<0.001	1.913	0.056
Retention of Urine—R33.9	2.621	0.678	3.866	**<0.001**

SE, Standard error of the coefficient; IV, Intravenous; RBC, Red Blood Cells.
Bold values are statistically significant values defined as p-value < 0.05.

system which acts to inhibit the detrusor muscle (27). Medications such as opioids are known to play dual functions by inhibiting parasympathetic and stimulating sympathetic innervations (28).

Thus far, no highly reliable and easily available prediction tools have been developed to identify *a priori* who is at increased risk for its development. Here, we present a model leveraging machine learning to classify the risk of a patient developing POUR following lumbar spine surgery using patient, surgical and anesthesia characteristics. Using machine learning, we were able to condense more than 90 variables associated with POUR in univariate analysis to a 14-variable logistic regression model and eventually constructed an eleven-node decision tree after hyper-parametric tuning of selected predictors from the LASSO regression model, with a final accuracy for the decision tree model of 87.8% on a confusion matrix and AUC-ROC of 0.737. This accuracy outperforms all previously available models and hence offers a novel and improved predictive tool for POUR.

The incidence of POUR within our study was 26% and is well within the incidence of POUR (5.6%–38%) reported across diverse studies of lumbar spine surgery (3, 5–7, 9, 12, 13). Previous studies have contained extensive inclusion and exclusion criteria for their models of POUR. We chose to include a heterogeneous patient population within our analysis to better understand how we can comprehensively evaluate the lumbar spine surgery population for the

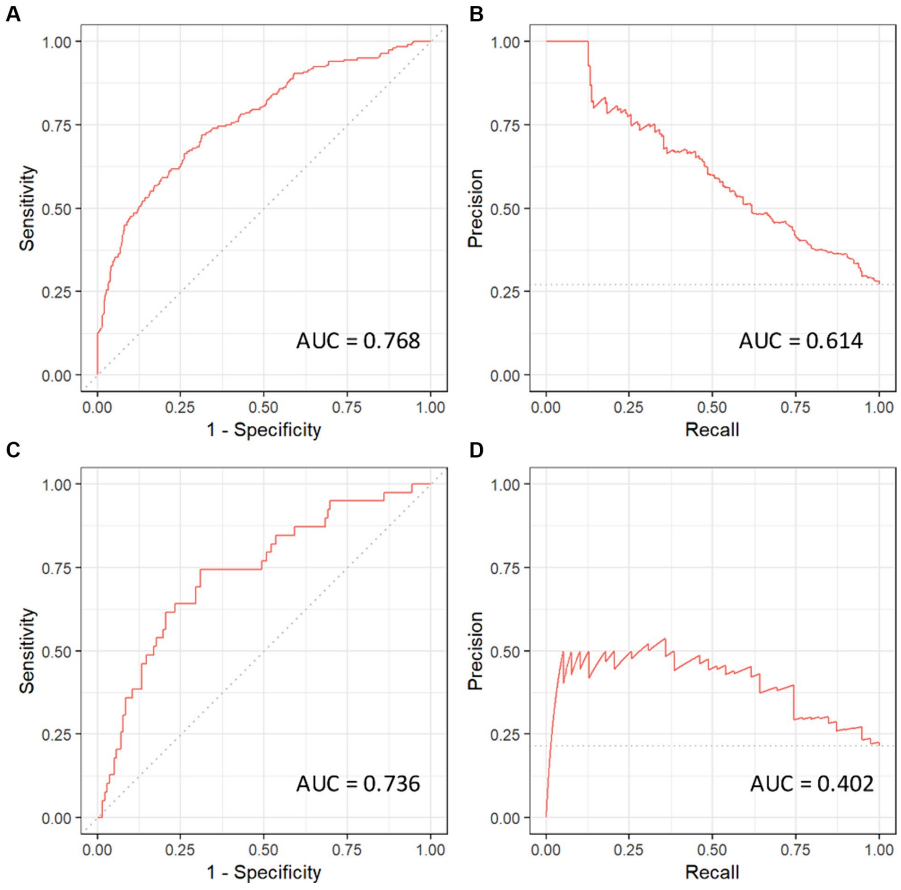


FIGURE 4
Receiver operating curves (ROC) and precision recall curves (PRC) for the logistic regression model. AUC, area under curve. **(A)** ROC-AUC for training set. **(B)** PRC-AUC for training set. **(C)** ROC-AUC for testing set. **(D)** PRC-AUC for testing set.

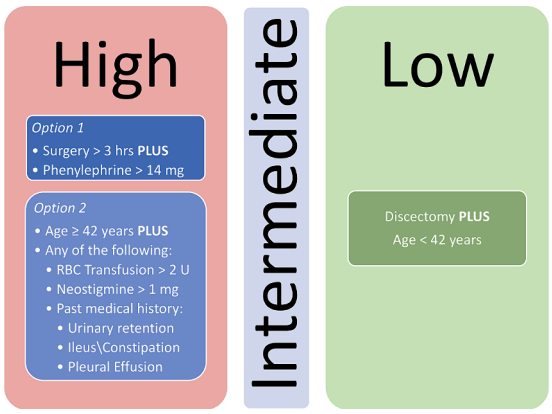


FIGURE 5
Decision tree model for stratifying patients for risk of POUR following lumbar spine surgery. RBC, Red blood cell.

development of POUR. By utilizing the logistic and LASSO regression models, a decision tree was able to be constructed that outperforms any prior predictive tool with accuracy of 0.878.

5.1 Limitations and future aims

Our model has limitations. As with all algorithms, it is only as accurate as the data which it contains. In this case, it is derived from a large tertiary care referral center where comprehensive data about a patient's past medical and surgical history may not be complete. We minimized this variability by extracting the medical history of patients from Epic's Care Everywhere network (Epic Systems Corporation) which accesses patient's medical charts from hundreds of other healthcare organizations, not exclusively our hospitals electronic medical record. Likewise, the study was retrospectively designed which carries biases inherent to a retrospective study.

This study was aimed at prediction of POUR, and not at interpretation of component variables. It serves as a diagnostic tool for POUR instead of identifying the critical variables that cause it. It can be tempting to elaborate on the meaning of predictors featured in the final model; however, these specific predictors are likely confounded by extensive patient and surgical variables and would warrant further prospective investigation. For factors such as phenylephrine (used for intraoperative blood pressure augmentation), a feasible alternative that is not associated with POUR, regardless of causality between the factor and POUR, might not exist. However, the use of intraoperative urinary catheters which appears to be statistically significant in all models, presents a potentially modifiable variable. While this variable is extensively confounded by surgical time and associated anesthesia requirements via medications and fluids, it remains important to investigate. Additionally, further improvement in the predictive capabilities of this model can be achieved by including baseline bladder/urologic functional status and preoperative urologic medication requirements.

6 Conclusion

In conclusion, we describe a highly accurate postoperative predictive model for POUR following lumbar spine using diverse preoperative and operative (surgical and anesthesia) variables. We were able to leverage machine learning to develop a 14 variable logistic regression model with an ROC-AUC of 0.737 and a decision tree model with an accuracy of 87.8%. These models substantially outperform previously published models of POUR in this patient population and include a greater spectrum of variables to highlight the effect of many less frequently appreciated variables. Furthermore, the final decision tree model is easy to implement clinically and can be put forth toward further studies aimed at preventing POUR following lumbar spine surgery. A prospective, multi-center study is needed to further validate our prediction model.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the University of Florida Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

SM: Methodology, Writing – original draft. KP: Conceptualization, Writing – review & editing. YM: Data curation, Resources, Writing – original draft. SY: Formal analysis, Visualization, Writing – original draft. CM: Conceptualization, Writing – review & editing. BL-W: Conceptualization, Formal analysis, Writing – review & editing. SR: Supervision, Writing – review & editing. MD: Project administration, Supervision, Writing – review & editing. KB: Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Baldini G, Bagry H, Aprikian A, Carli F, Warner DS, Warner MA. Postoperative urinary retention: anesthetic and perioperative considerations. *Anesthesiology*. (2009) 110:1139–57. doi: 10.1097/ALN.0b013e31819f7aea

2. Pomajzl AJ, Siref LE. Post-op urinary retention. In: StatPearls. Treasure Island (FL): StatPearls publishing (2022). Available at: http://www.ncbi.nlm.nih.gov/books/NBK549844/ (Accessed 18 February 2022).

3. Boulis NM, Mian FS, Rodriguez D, Cho E, Hoff JT. Urinary retention following routine neurosurgical spine procedures. *Surg Neurol*. (2001) 55:23–7. doi: 10.1016/S0090-3019(01)00331-7

4. Cremins M, Vellanky S, McCann G, Mancini M, Sanzari L, Yannopoulos A. Considering healthcare value and associated risk factors with postoperative urinary retention after elective laminectomy. *Spine J*. (2020) 20:701–7. doi: 10.1016/j.spinee.2020.01.012

5. Gandhi SD, Patel SA, Maltenfort M, Anderson DG, Vaccaro AR, Albert TJ, et al. Patient and surgical factors associated with postoperative urinary retention after lumbar spine surgery. *Spine*. (2014) 39:1905–9. doi: 10.1097/BRS.0000000000000572

6. Jellish WS, Thalji Z, Stevenson K, Shea J. A prospective randomized study comparing short-and intermediate-term perioperative outcome variables after spinal or general anesthesia for lumbar disk and laminectomy surgery. *Anesth Analg*. (1996) 83:559–64. doi: 10.1213/00000539-199609000-00021

7. Kopel J, Sharma DP. Spinal surgery and urinary retention: a review of the literature. *J Clin Urol*. (2020) 14:4–9. doi: 10.1177/2051415820916932

8. Porche K, Samra R, Melnick K, Brennan M, Vaziri S, Seubert C, et al. Enhanced recovery after surgery (ERAS) for open transforaminal lumbar interbody fusion: a retrospective propensity-matched cohort study. *Spine J*. (2021) 22:399–410. doi: 10.1016/j.spinee.2021.10.007

9. Golubovsky JL, Ilyas H, Chen J, Tanenbaum JE, Mroz TE, Steinmetz MP. Risk factors and associated complications for postoperative urinary retention after lumbar surgery for lumbar spinal stenosis. *Spine J*. (2018) 18:1533–9. doi: 10.1016/j.spinee.2018.01.022

10. Martinez OV, Civetta JM, Anderson K, Roger S, Murtha M, Malinin TI, et al. Bacteriuria in the catheterized surgical intensive care patient. *Crit Care Med*. (1986) 14:188–91. doi: 10.1097/00003246-198603000-00003

11. Zakaria HM, Lipphardt M, Bazydlo M, Xiao S, Schultz L, Chedid M, et al. The preoperative risks and two-year sequelae of postoperative urinary retention: analysis of the Michigan spine surgery improvement collaborative (MSSIC). *World Neurosurg*. (2020) 133:e619–26. doi: 10.1016/j.wneu.2019.09.107

12. Altschul D, Kobets A, Nakhla J, Jada A, Nasser R, Kinon MD, et al. Postoperative urinary retention in patients undergoing elective spinal surgery. *J Neurosurg Spine*. (2017) 26:229–34. doi: 10.3171/2016.8.SPINE151371

13. Lee S, Kim CH, Chung CK, Park SB, Yang SH, Kim SH, et al. Risk factor analysis for postoperative urinary retention after surgery for degenerative lumbar spinal stenosis. *Spine J*. (2017) 17:469–77. doi: 10.1016/j.spinee.2016.03.017

14. DelSole EM, Keck WL, Patel AA. The state of machine learning in spine surgery: a systematic review. *Clin Spine Surg*. (2022) 35:80–9. doi: 10.1097/BSD.0000000000001208

15. Porche K, Maciel CB, Lucke-Wold B, Robicsek SA, Chalouhi N, Brennan M, et al. Preoperative prediction of postoperative urinary retention in lumbar surgery: a comparison of regression to multilayer neural network. *J Neurosurg Spine*. (2021) 36:32–41. doi: 10.3171/2021.3.SPINE21189

16. Sfeir S, Mansour N. Post operative analgesia with intrathecal morphine. *Middle East J Anaesthesiol*. (2005) 18:133–9.

17. Mormol JD, Basques BA, Harada GK, Louie PK, Alter K, Goldberg E, et al. Risk factors associated with development of urinary retention following posterior lumbar spinal fusion: special attention to the use of Glycopyrrolate in anesthesia reversal. *Spine*. (2021) 46:E133–8. doi: 10.1097/BRS.0000000000003678

18. Bowman JJ, Edwards CC 2nd, Dean C, Park J, Edwards CC Sr. Incidence and risk factors for postoperative urinary retention following lumbar spine fusion. *Clin Spine Surg*. (2021) 34:E397–402. doi: 10.1097/BSD.0000000000001202

19. Madani AH, Aval HB, Mokhtari G, Nasseh H, Esmaeili S, Shakiba M, et al. Effectiveness of tamsulosin in prevention of post-operative urinary retention: a randomized double-blind placebo-controlled study. *Int Braz J Urol Off J Braz Soc Urol*. (2014) 40:30–6. doi: 10.1590/S1677-5538.IBJU.2014.01.05

20. Aiyer SN, Kumar A, Shetty AP, Kanna RM, Rajasekaran S. Factors influencing postoperative urinary retention following elective posterior lumbar spine surgery: a prospective study. *Asian Spine J*. (2018) 12:1100–5. doi: 10.31616/asj.2018.12.6.1100

21. Knight BA, Bayne AP, Zusman N, Barney N, Yang S. Postoperative management factors affect urinary retention following posterior spinal fusion for adolescent idiopathic scoliosis. *Spine Deform*. (2020) 8:703–9. doi: 10.1007/s43390-020-00090-9

22. Chang Y, Chi K-Y, Tai T-W, Cheng YS, Lee PH, Huang CC, et al. Risk factors for postoperative urinary retention following elective spine surgery: a meta-analysis. *Spine J*. (2021) 21:1802–11. doi: 10.1016/j.spinee.2021.05.009

23. Petros JG, Bradley TM. Factors influencing postoperative urinary retention in patients undergoing surgery for benign anorectal disease. *Am J Surg*. (1990) 159:374–6. doi: 10.1016/S0002-9610(05)81274-7

24. Faas CL, Acosta FJ, Campbell MDR, O'Hagan CE, Newton SE, Zagalaniczny K. The effects of spinal anesthesia vs epidural anesthesia on 3 potential postoperative complications: pain, urinary retention, and mobility following inguinal herniorrhaphy. *AANA J*. (2002) 70:441–7.

25. Brier GW. Verification of Forecasts Expressed in Terms of Probability. *Mon Weather Rev*. (1950) 78:1–3. doi: 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

26. Kamphuis ET, Ionescu TI, Kuipers PW, de Gier J, van Venrooij GE, Boon TA. Recovery of storage and emptying functions of the urinary bladder after spinal anesthesia with lidocaine and with bupivacaine in men. *Anesthesiology*. (1998) 88:310–6. doi: 10.1097/00000542-199802000-00007

27. Toyonaga T, Matsushima M, Sogawa N, Jiang SF, Matsumura N, Shimojima Y, et al. Postoperative urinary retention after surgery for benign anorectal disease: potential risk factors and strategy for prevention. *Int J Color Dis*. (2006) 21:676–82. doi: 10.1007/s00384-005-0077-2

28. Elsamra SE, Ellsworth P. Effects of analgesic and anesthetic medications on lower urinary tract function. *Urol Nurs*. (2012) 32:60–8. doi: 10.7257/1053-816X.2012.32.2.60

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 03 October 2024
DOI 10.3389/fneur.2024.1419608

Check for updates

OPEN ACCESS

EDITED BY
Alejandro Rabinstein,
Mayo Clinic, United States

REVIEWED BY
Ping Hu,
Second Affiliated Hospital of Nanchang University, China
Muhannad Seyam,
University of Vermont, United States

*CORRESPONDENCE
Xinjiang Yan
✉ 1582344125@qq.com

RECEIVED 10 June 2024
ACCEPTED 18 September 2024
PUBLISHED 03 October 2024

CITATION
Xu J, Yuan C, Yu G, Li H, Dong Q, Mao D, Zhan C and Yan X (2024) Predicting cerebral edema in patients with spontaneous intracerebral hemorrhage using machine learning.
Front. Neurol. 15:1419608.
doi: 10.3389/fneur.2024.1419608

COPYRIGHT
© 2024 Xu, Yuan, Yu, Li, Dong, Mao, Zhan and Yan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Predicting cerebral edema in patients with spontaneous intracerebral hemorrhage using machine learning

Jiangbao Xu[1], Cuijie Yuan[1], Guofeng Yu[1], Hao Li[2,3], Qiutong Dong[2,3], Dandan Mao[1], Chengpeng Zhan[1] and Xinjiang Yan[1]*

[1]The Quzhou Affiliated Hospital of Wenzhou Medical University, Quzhou People's Hospital, Quzhou, China, [2]Postgraduate Training Base Alliance of Wenzhou Medical University, Wenzhou, China, [3]Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, China

Background: The early prediction of cerebral edema changes in patients with spontaneous intracerebral hemorrhage (SICH) may facilitate earlier interventions and result in improved outcomes. This study aimed to develop and validate machine learning models to predict cerebral edema changes within 72 h, using readily available clinical parameters, and to identify relevant influencing factors.

Methods: An observational study was conducted between April 2021 and October 2023 at the Quzhou Affiliated Hospital of Wenzhou Medical University. After preprocessing the data, the study population was randomly divided into training and internal validation cohorts in a 7:3 ratio (training: $N = 150$; validation: $N = 65$). The most relevant variables were selected using Support Vector Machine Recursive Feature Elimination (SVM-RFE) and Least Absolute Shrinkage and Selection Operator (LASSO) algorithms. The predictive performance of random forest (RF), GDBT, linear regression (LR), and XGBoost models was evaluated using the area under the receiver operating characteristic curve (AUROC), precision–recall curve (AUPRC), accuracy, F1-score, precision, recall, sensitivity, and specificity. Feature importance was calculated, and the SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) methods were employed to explain the top-performing model.

Results: A total of 84 (39.1%) patients developed cerebral edema changes. In the validation cohort, GDBT outperformed LR and RF, achieving an AUC of 0.654 (95% CI: 0.611–0.699) compared to LR of 0.578 (95% CI, 0.535–0.623, DeLong: $p = 0.197$) and RF of 0.624 (95% CI, 0.588–0.687, DeLong: $p = 0.236$). XGBoost also demonstrated similar performance with an AUC of 0.660 (95% CI, 0.611–0.711, DeLong: $p = 0.963$). However, in the training set, GDBT still outperformed XGBoost, with an AUC of 0.603 ± 0.100 compared to XGBoost of 0.575 ± 0.096. SHAP analysis revealed that serum sodium, HDL, subarachnoid hemorrhage volume, sex, and left basal ganglia hemorrhage volume were the top five most important features for predicting cerebral edema changes in the GDBT model.

Conclusion: The GDBT model demonstrated the best performance in predicting 72-h changes in cerebral edema. It has the potential to assist clinicians in identifying high-risk patients and guiding clinical decision-making.

KEYWORDS

SICH, cerebral edema, random forest, GDBT, XGBoost

Introduction

Spontaneous intracerebral hemorrhage (SICH) is a prevalent subtype of stroke, with a mortality rate significantly higher than ischemic stroke. Approximately 20–30% of SICH patients die within 3 months (1–3). The high incidence and mortality rates pose a significant threat to public health (4, 5). Cerebral edema, a common complication of SICH, involves the accumulation of excess water in the brain tissues adjacent to the hemorrhage. This can lead to severe consequences, including compromised blood flow, intracranial pressure shifts, and neuronal damage (6, 7). Timely identification of edema development and its influencing factors is crucial for optimizing patient care, allocating resources effectively, and reducing healthcare costs. Cerebral edema typically appears within 24–72 h after bleeding, peaks 2–7 days later, and can persist for up to 2 weeks. To monitor edema progression, patients with SICH undergo head CT scans at admission, 24 and 72 h post-admission. Subsequent scans may be ordered based on clinical changes. This study aimed to develop and validate machine learning models capable of predicting changes in cerebral edema within the first 72 h following SICH (8).

Over the past few years, advances in imaging omics have refined the use of CT scans for evaluating brain edema (9). In addition, machine learning algorithms have demonstrated significant promise in predicting medical outcomes and complications, aiding clinicians in making informed decisions and enhancing patient care (10–13). These advancements inform the development of accurate and reliable prognostic models to identify patients at risk of severe cerebral edema, enabling healthcare providers to implement targeted preventive strategies and interventions.

This study sought to develop and validate a machine learning-based prognostic model that could evaluate the progression patterns and influencing factors of cerebral edema, considering various patient attributes and clinical determinants. We aimed to compare the predictive accuracy and clinical utility of different machine learning algorithms. Ultimately, our goal was to provide clinicians with valuable tools for the early identification of patients at risk of severe cerebral edema, enabling the implementation of targeted preventive strategies to reduce its prevalence.

Methods

This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Data source

This observational study was conducted between April 2021 and October 2023 at the Quzhou Affiliated Hospital of Wenzhou Medical University. The study protocol was approved by the hospital's ethical board (reference number: LW2023-163) and adhered to the principles of the Declaration of Helsinki. All patients provided informed consent through their relatives, and no patient data were used in a way that could pose a risk to them. Patients with SICH were included if they met the following criteria: (1) admission within 72 h after first-ever stroke; (2) SICH confirmed by head computerized tomography (CT) scan; (3) hospitalization within 24 h after the onset of stroke symptom; and (4) age of 18 years or greater. Exclusion criteria encompassed the following:(1) secondary brain bleeding as a result of congenital or acquired coagulation abnormalities, hemorrhagic transformation of cerebral infarction, moyamoya disease, cerebral aneurysm, and arteriovenous malformation or tumor; (2) primary intraventricular bleeding; (3) presence of previous neurological diseases, such as brain tumors and severe head trauma; and (4) coexistence with severe systemic diseases, for example, malignancies, immune deficiency syndromes, and severe heart, liver, lung, or kidney dysfunction.

A total of 215 patients presented to the emergency department with suspected SICH, which was confirmed by head CT scans. All CT scans were conducted following the radiology department's protocol by radiologists blinded to clinical information. To ensure data relevance, we collected 51 basic patient characteristics at admission (detailed in Supporting Material 1). In addition, we gathered 38 imaging characteristics, including clinical review times at 6, 24, 72 h, and subsequent hours until the absence of hematoma and edema was confirmed by two senior doctors (Supporting Material 1).

The cerebral edema volume was calculated using two methods: (1) Image-based analysis: The boundaries of hematoma and edema were delineated on CT scans using image browser measurement software. The hematoma volume was calculated by summing the areas of each layer, while the edema volume was determined by subtracting the hematoma volume from the combined volume of hematoma and surrounding brain edema. (2) Formula-based analysis (verification): The hematoma length and width in the maximum plane were measured as A and B, respectively, and the thickness (number of layers, C) was calculated. The hematoma volume was calculated as 1/2 ABC, with layers categorized as 75% (layer 1), 75–25% (1/2 layer), and < 25% (excluded). The edema volume was determined as the difference between the combined volume of blood and edema and the hematoma volume. Absolute hematoma and edema volumes were assessed at each time point. The primary outcome measure was defined as an increase in cerebral edema volume between baseline and repeat imaging of more than 6 mL or a relative increase of >33% within 72 h (14–17). To optimize statistical power and minimize bias, multiple imputation using random forests was employed to supplement missing values. The imputed data were then randomly stratified into training ($N = 150$) and validation cohorts ($N = 65$) in a 7:3 ratio.

Feature selection

To prevent variable misselection, we employed a rigorous variable selection approach using a training cohort to identify the most relevant predictors for constructing a predictive model. Initially, pairwise Pearson's correlation matrices were used to assess the collinearity of clinical variables. Collinearity occurs when two or more predictors exhibit a strong correlation ($r > 0.8$), complicating the evaluation of each variable's unique contribution to the outcome. Therefore, we have chosen to remove the more readily available variables from the collinear variables. Subsequently, we used the Minimum Absolute Shrinkage and Selection Operator

(LASSO) and the SVM-RFE algorithm in a two-step process. LASSO is a regularization technique that performs variable selection and coefficient estimation by applying constraints to the sum of the absolute values of the model parameters. This process causes some of the coefficients to be narrowed down to zero, effectively excluding them from the final model. Then, the SVM-RFE algorithm was used for further variable selection. The SVM-RFE algorithm enables the machine learning algorithm to continuously reduce the number of features, verify the performance of the model, and finally achieve the optimal number of features for screening. By using the SVM-RFE algorithm, we obtain another important set of predictors. Finally, the intersection of predictors determined by the LASSO and SVM-RFE algorithms is employed to ensure that only the most relevant and robust variables are included in the development of our predictive models. This combined approach aims to improve the accuracy and generalizability of the model while reducing the risk of overfitting or including irrelevant predictors.

Model development and validation

We employed four machine learning classifiers—extreme gradient boosting (XGBoost), random forest (RF), linear regression (LR), and gradient-boosted decision trees (GDBT) (18–21)—to develop predictive models for the risk of 72-h brain edema growth. All models incorporated the same input variables. Grid and random hyperparameter searches were conducted on the training data to identify optimal hyperparameters for each model, with performance evaluated using the area under the receiver operating characteristic curve (AUROC), precision–recall curve (AUPRC), F1 score, precision, recall, sensitivity, and specificity. To interpret the best-performing model, Shapley Additive exPlanations (SHAP) (22) and Local Interpretable Model-Agnostic Explanations (LIME) (23) were applied to provide consistent and locally accurate variable importance values, enhancing our understanding of the model's predictive capabilities.

Dataset selection

Given the limited sample size in this study, which can introduce bias, a 5-fold cross-validation was employed to ensure objectivity and minimize sampling bias. To select the dataset with the greatest statistical significance, a Wilcoxon rank-sum test was performed. The dataset with the largest *p*-value was chosen as illustrated in Figure 1.

Tuning of hyperparameters

XGBoost

XGBoost, a widely used and powerful ensemble technique, is based on the gradient boosting framework. It combines the predictions of multiple weak learners, primarily decision trees, to create a more accurate and robust model. XGBoost implements machine learning algorithms within the Gradient Boosting framework. The optimal parameters were determined using the "xgboost" package and 5-fold cross-validation, as illustrated in Figure 2.

Random forest

Random forest algorithms utilize tree-based models, combining multiple decision trees through bootstrapping to improve predictive accuracy (19). The optimal number of trees was determined using 5-fold cross-validation with the "randomForest" package, as illustrated in Figure 2.

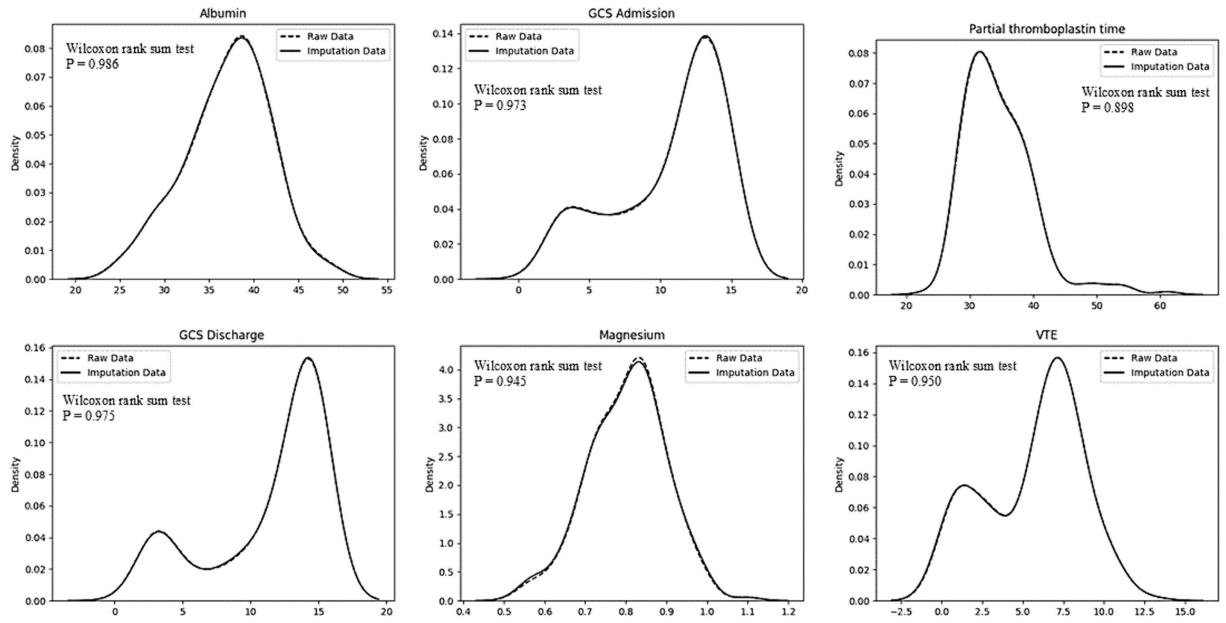


FIGURE 1
Dataset selection plots.

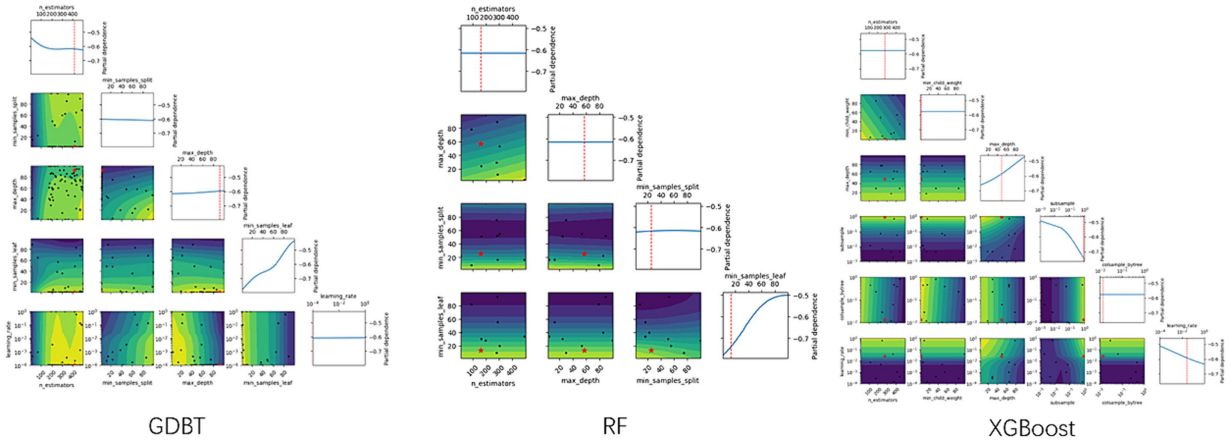


FIGURE 2
Hyperparameter selection plots.

Linear regression

Linear regression is a widely used statistical method for modeling binary outcomes. The most common approach is least squares, which aims to minimize the average squared error between predicted and observed values. To select optimal variables and construct an LR model, we employed backward stepwise regression based on the Akaike information criterion. The "MASS" package in R software was utilized to fit the model (24).

Gradient-boosted decision tree (GBDT)

The GBDT model iteratively calculates residuals at each step and establishes the model by moving in the negative gradient direction of these residuals. GBDT's powerful, flexible, efficient, and accurate predictive capabilities have made it a popular machine-learning algorithm for analyzing and processing abstract data. Optimal parameters for the GBDT model were determined using the "GBDT" package and 5-fold cross-validation, as illustrated in Figure 2.

Statistical analysis

Prior to formal analysis, the Kolmogorov–Smirnov test was used to assess data distribution. Continuous variables were analyzed using either the independent *t*-test (for normally distributed data) or the Mann–Whitney *U*-test (for non-normally distributed data) and were presented as mean ± standard deviation (SD) or median with interquartile range (IQR), respectively. Categorical variables were analyzed using the chi-square test for large samples or Fisher's exact test for small samples and are expressed as frequencies (percentages). To compare the area under the curve (AUC) of the different models statistically, the DeLong test was used. All statistical tests were two-tailed, and a $p < 0.05$ was considered statistically significant. In addition, the study adhered to the rule of thumb of having at least 10 events per variable for robust analysis. Statistical analyses were performed using R (version 4.2.2; R Foundation for Statistical Computing) and Python (version 3.9.0; Python Software Foundation).

Results

Patient characteristics

The dataset comprised information on 215 patients with SICH, including 949 imaging CT scans. The total number of CT scans was determined at specific time frames: 6, 24, 72 h, and subsequent hours until the absence of hematoma and edema was confirmed by two senior doctors. Of these patients, 86 (40%) exhibited cerebral edema expansion (edema volume increased by more than 6 mL or by >33% relative to the last measurement) within 72 h. The cohort included 143 male (66.5%) and 72 female (33.5%) patients, with 60 male (69.8%) and 26 female (30.2%) patients experiencing dilated cerebral edema. No significant differences were observed in baseline characteristics between the training and validation groups. Tables 1, 2 provide detailed baseline patient characteristics.

Feature selection

As shown in Figure 3, no pairwise Pearson's correlations between continuous variables exceeded 0.8, indicating the absence of collinearity. Consequently, all variables were included in the subsequent feature selection process. SVM-RFE identified 31 important predictors (Supporting Material 1), while the LASSO regression algorithm selected 26 (Supporting Material 1). Ultimately, 20 factors emerged as significant predictors of the outcome (Figure 4), including sex, diabetes history, hypertension history, alcohol history, ventricular drainage, hemostatic treatment, decompressive craniectomy, antihypertensive treatment, antiemesis and antacid, HDL, cholesterol, alanine aminotransferase, serum magnesium, serum sodium, CRP, admission Barthel ADL Index, cerebral subarachnoid hemorrhage volume, subdural hemorrhage volume, and hemorrhage in the left cerebellum, left basal ganglia, or left parietal lobe. These selected features were integrated into four machine learning classifiers—GDBT, LR, RF, and XGBoost—to develop the predictive model.

TABLE 1 Summary table of categorized data.

Characteristics	Totals (n = 215)	Training cohort (n = 150)	Validation cohort (n = 65)	P-value
Sex	143 (0.665)	101 (0.673)	42 (0.646)	0.818
History of stroke	58 (0.270)	40 (0.267)	18 (0.277)	1.000
History of diabetes	29 (0135)	18 (0.120)	11 (0.169)	0.451
History of atrial fibrillation	12 (0.056)	8 (0.053)	4 (0.062)	1.000
History of coronary heart disease	92 (0.428)	71 (0.473)	21 (0.323)	0.058
History of hypertension	183 (0.851)	124 (0.827)	59 (0.908)	0.185
History of smoking	69 (0.321)	53 (0.353)	16 (0.246)	0.165
History of alcohol consumption	73 (0.340)	53 (0.353)	20 (0.308)	0.623
History of hyperlipidemia	4 (0.019)	1 (0.007)	3 (0.046)	0.156
Cranial decompression Hematoma removal	67 (0.312)	43 (0.287)	24 (0.369)	0.298
Ventricular drainage	71 (0.330)	44 (0.293)	27 (0.415)	0.112
Hemostatic therapy	200 (0.930)	141 (0.940)	59 (0.908)	0.574
Cranial pressure-lowering therapy	194 (0.902)	138 (0.920)	56 (0.862)	0.282
Antihypertensive treatment	208 (0.967)	146 (0.973)	62 (0.954)	0.748
Antiemetic and antacid	211 (0.981)	148 (0.987)	63 (0.969)	0.749
Lipid-lowering therapy	20 (0.093)	17 (0.113)	3 (0.046)	0.193

Tuning of hyperparameters

As shown in Table 3 and Figure 2, the optimal hyperparameters for the GDBT models were as follows: n_estimators (10–250), min_samples_split (2–25), max_features (0.1–0.999), max_depth (3–15), min_samples_leaf (1–25), and learning_rate (0.001–0.3). For the extreme gradient boosting (XGBoost) models, the optimal hyperparameters were as follows: n_estimators (10–250), min_samples_split (1–25), max_depth (3–15), subsample (0.001–1), colsample_bytree (0.01–1), and learning_rate (0.001–0.3). Finally, the optimal hyperparameters for the RF models were as follows: n_estimators (10–250), max_depth (3–15), min_samples_split (2–25), min_samples_leaf (1–25), and max_features (0.1–0.999).

Development and validation of prediction models

When evaluating model performance on the validation cohort, our results demonstrated that the GDBT model, with an AUC value of 0.654 (95% CI: 0.611–0.699), outperformed the LR and RF models, which yielded AUC values of 0.578 (95% CI: 0.535–0.623, DeLong: $p = 0.197$) and 0.624 (95% CI: 0.588–0.687, DeLong: $p = 0.236$), respectively. Similarly, GDBT outperformed XGBoost, with an AUC of 0.660 (95% CI: 0.611–0.711, DeLong: $p = 0.963$). However, in the training set, GDBT (AUC = 0.603 ± 0.100) outperformed XGBoost (AUC = 0.575 ± 0.096). To mitigate the effects of random sampling, we repeated this process 50 times. Over-validation revealed that the LR model exhibited overfitting, performing poorly on the independent dataset despite good performance on the training set. In contrast, the GDBT model demonstrated greater stability and superior performance in both the training and validation sets compared to XGBoost. Based on these results, we selected the GDBT model for subsequent experiments, as summarized in Table 4. Receiver operating curves and precision–recall curves for the models are depicted in Figure 5.

Model explainability

The SHAP summary plot (Figure 6) illustrates the relative importance of the 20 predictors in the GDBT model. We discovered that serum sodium, HDL cholesterol, subarachnoid hemorrhage volume, sex, and left basal ganglia hemorrhage volume were the five most significant features for predicting cerebral edema changes in the SHAP (GDBT) model.

The LIME interpreter was applied to data generated by the GDBT model to examine classification outcomes. Each case's feature weights are depicted in Figure 7, with green indicating factors favoring the outcome and red representing those opposing it. In case 1, the 100% predicted increase in edema was likely attributed to sex (male), cerebral subarachnoid hemorrhage volume within the range of 0.02–0.15 mL, alcohol use history, HDL levels between 1.2 and 1.44 mmol/L, serum sodium between 141.55 and 143.45 mmol/L, absence of subdural hemorrhage, alanine aminotransferase levels within 13.4–19.4 U/L, and left parietal lobe volume between 0.13 mL and 0.67 mL (favoring variables). However, no drinking history and no cerebral ventricular drainage were the opposite variables. Conversely, in case 2, the 99.8% prediction of no edema increase was likely due to serum sodium levels exceeding 143.45 mmol/L, left parietal bleeding volume within the 0.13–0.67 mL range, alanine aminotransferase levels between 19.04 U/L and 27.92 U/L, and CRP levels within 1.58–3.61 mg/L (favoring variables). Sex (male), HDL levels exceeding 1.44 mmol/L, and a subdural hemorrhage volume >0.1 mL were the opposite variables.

10.3389/fneur.2024.1419608

TABLE 2 Summary table of continuous variable data.

Characteristics	Totals (*n* = 215)	Training cohort (*n* = 150)	Validation cohort (*n* = 65)	*P*-value
Age	66.0 (56.0, 75.0)	66.5 (56.0, 76.75)	66.0 (56.0, 73.0)	0.531
Time from onset to first imaging	6.0 (4.0, 10.5)	5.5 (4.0, 10.0)	6.0 (4.0, 12.0)	0.637
Systolic blood pressure	158.0 (142.0, 174.0)	158.5 (142.25, 174.0)	158.0 (142.0, 173.0)	0.723
Diastolic blood pressure	88.0 (78.0, 97.0)	87.0 (77.25, 97.0)	88.0 (78.0, 100.0)	0.647
VTE	6.0 (3.0, 8.0)	7.0 (3.0, 8.0)	6.0 (3.0, 7.0)	0.351
Hemoglobin	133.0 (122.0, 143.0)	133.0 (123.0, 144.0)	133.0 (121.0, 141.0)	0.436
Blood platelet count	176.0 (128.0, 212.5)	177.5 (135.5, 210.75)	169.0 (117.0, 218.0)	0.563
Partial thromboplastin time	33.4 (30.65, 37.3)	33.85 (30.725, 38.15)	32.7 (30.4, 36.0)	0.057
Prothrombin time (blood clotting enzyme)	17.1 (16.3, 17.7)	17.1 (16.3, 17.6)	17.1 (16.4, 17.8)	0.604
Prothrombin time	13.2 (12.6, 13.85)	13.2 (12.6, 13.85)	13.2 (12.8, 13.8)	0.845
Triglyceride	0.97 (0.705, 1.45)	0.95 (0.712, 1.438)	1.05 (0.69, 1.49)	0.567
Cholesterol	4.03 (3.42, 4.72)	3.925 (3.43, 4.742)	4.24 (3.41, 4.63)	0.460
High-density lipoprotein (HDL) cholesterol	1.22 (0.99, 1.45)	1.2 (0.982, 1.402)	1.26 (1.03, 1.5)	0.280
Glutamic-pyruvic transaminase	18.4 (13.0, 28.55)	18.9 (13.125, 28.575)	17.6 (12.6, 28.3)	0.963
Creatine kinase	106.3 (69.15, 184.05)	107.5 (69.6, 181.825)	103.3 (64.7, 188.7)	0.571
Lactate dehydrogenase	212.2 (184.75, 255.25)	209.2 (186.175, 250.85)	220.7 (183.8, 277.6)	0.271
Creatine kinase isoenzyme	20.8 (15.65, 27.55)	20.95 (16.2, 26.875)	20.1 (15.1, 29.5)	0.652
Magnesium	0.81 (0.74, 0.86)	0.805 (0.732, 0.86)	0.82 (0.75, 0.88)	0.232
Sugar	6.39 (5.31, 8.39)	6.375 (5.262, 8.278)	6.94 (5.41, 8.6)	0.356
Calcium	2.19 (2.1, 2.26)	2.19 (2.1, 2.278)	2.18 (2.1, 2.24)	0.629
Sodium	141.0 (139.15, 143.2)	141.1 (138.825, 143.1)	141.0 (139.9, 143.3)	0.310
Albumin	37.7 (34.25, 40.3)	37.95 (34.3, 40.6)	37.0 (34.1, 39.8)	0.387
C-reactive protein	4.26 (1.73, 17.005)	4.175 (1.762, 13.252)	4.35 (1.46, 19.2)	0.737
ADL admission	10.0 (0.0, 35.0)	10.0 (0.0, 35.0)	10.0 (0.0, 35.0)	0.808
ADL discharge	40.0 (0.0, 60.0)	37.5(0.0, 60.0)	50.0(0.0, 70.0)	0.172
Admission mRS score	4.0 (3.0, 5.0)	4.0 (4.0, 5.0)	4.0 (3.0, 5.0)	0.174
Mean_median_offset	0.0 (−0.517, 1.224)	0.0 (−0.375, 1.192)	0.0 (−0.533, 1.214)	0.957
Mean_cerebral subarachnoid hemorrhage volume	0.014 (0.0, 0.194)	0.018 (0.0, 0.167)	0.0 (0.0, 0.267)	0.563
Mean_subdural hemorrhage.3. Epidural hemorrhage4	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.125)	0.250
Mean_subdural hemorrhage	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.809
Mean_CT5	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.760
Mean_brainstem	0.0 (0.0, 0.4)	0.0 (0.0, 0.333)	0.0 (0.0, 0.5)	0.777
Mean_left_cerebellum	0.0 (0.0, 0.167)	0.0 (0.0, 0.125)	0.0 (0.0, 0.25)	0.413
Mean_left_basal_ganglia	0.25 (0.0, 1.0)	0.225 (0.0, 1.0)	0.25 (0.0, 1.0)	0.793
Mean_left_frontal_lobe	0.0 (0.0, 0.354)	0.0 (0.0, 0.333)	0.0 (0.0, 0.375)	0.861
Mean_left_temporal_lobe	0.2 (0.0, 0.854)	0.167 (0.0, 0.75)	0.25 (0.0, 1.0)	0.463
Mean_left_thalamus	0.0 (0.0, 0.388)	0.0 (0.0, 0.333)	0.0 (0.0, 0.5)	0.677
Mean_left_parietal	0.0(0.0, 0.6)	0.125(0.0, 0.6)	0.0(0.0, 0.75)	0.540
Mean_left_occipital_lobe	0.0 (0.0, 0.062)	0.0 (0.0, 0.0)	0.0 (0.0, 0.143)	0.245
Mean_right_cerebellum	0.0 (0.0, 0.167)	0.0 (0.0, 0.2)	0.0 (0.0, 0.125)	0.701

(Continued)

TABLE 2 (Continued)

Characteristics	Totals (*n* = 215)	Training cohort (*n* = 150)	Validation cohort (*n* = 65)	*P*-value
Mean_right_basal_ganglia	0.2 (0.0, 1.0)	0.25 (0.0, 1.0)	0.0 (0.0, 1.0)	0.527
Mean_right_frontal_lobe	0.0 (0.0, 0.333)	0.0 (0.0, 0.333)	0.0 (0.0, 0.25)	0.843
Mean_right_temporal_lobe	0.0 (0.0, 0.667)	0.0 (0.0, 0.667)	0.0 (0.0, 0.5)	0.321
Mean_right_thalamus	0.0 (0.0, 0.5)	0.0 (0.0, 0.575)	0.0 (0.0, 0.333)	0.584
Mean_right_parietal	0.0 (0.0, 0.5)	0.0 (0.0, 0.5)	0.0 (0.0, 0.25)	0.208
Mean_right_occipital_lobe	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.514
Mean_fisher	0.5 (0.0, 1.31)	0.586 (0.0, 1.333)	0.429 (0.0, 1.167)	0.463

FIGURE 3
Pearson's correlation matrix thermodynamic chart.

Discussion

Hemorrhagic stroke, also known as cerebral hemorrhage, occurs when non-traumatic blood vessels in the brain rupture, leading to blood accumulation in the brain parenchyma. This condition constitutes 10–15% of all stroke cases and is characterized by rapid progression, severe neurological dysfunction, and a high mortality rate, particularly mortality rate during the acute phase (up to 100%). Increased intracranial pressure and cerebral herniation due to cerebral edema are major causes of death. Patients may also experience long-term neurological deficits, impacting their self-care ability and imposing substantial economic burdens on society and families. Early diagnosis and timely treatment are crucial for reducing mortality rates. To predict 72-h brain edema growth, we developed and validated

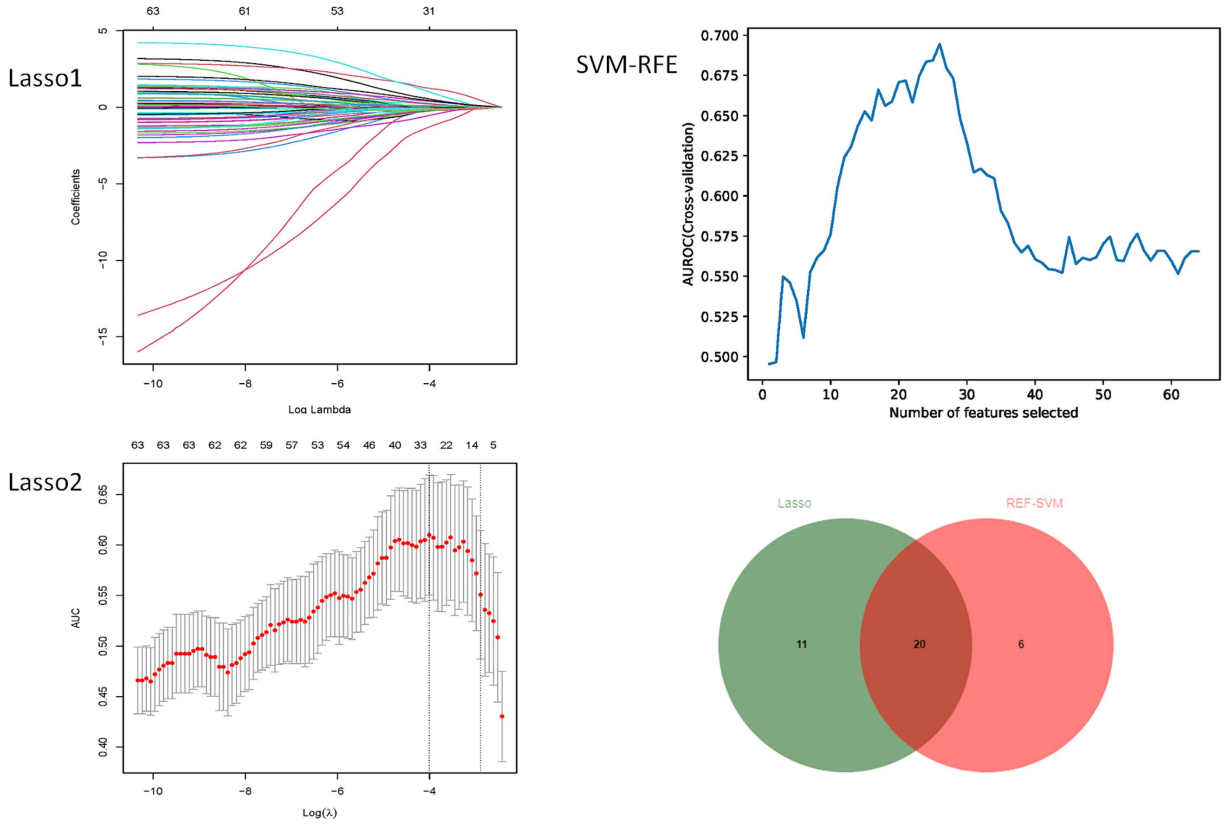


FIGURE 4
Variable screening diagram.

TABLE 3 Summary table of the model parameters.

GDBT		XGBoost		RF	
n_estimators	(10, 250)	n_estimators	(10, 250)	n_estimators	(10, 250)
min_samples_split	(2, 25)	min_child_weight	(1, 25)	max_depth	(3, 15)
max_features	(0.1, 0.999)	max_depth	(3, 15)	min_samples_split	(2, 25)
max_depth	(3, 15)	subsample	(0.001, 1)	min_samples_leaf	(1, 25)
min_samples_leaf	(1, 25)	colsample_bytree	(0.01, 1)	max_features	(0.1, 0.999)
learning_rate	(0.001,0.3)	learning_rate	(0.001,0.3)		

machine-learning models using four different algorithms (GDBT, LR, RF, and XGBoost). Twenty key predictors were identified, and internal and external validation demonstrated the superior performance and clinical applicability of the GDBT model (25–28).

Our importance analysis identified serum sodium levels as the most significant predictor of 72-h brain edema growth risk, aligning with previous research (29–33). These findings support the role of serum sodium as a valuable prognostic indicator in brain edema. Previous studies suggest that edema around hematoma is predominantly vasogenic in the early stages, transitioning to cytotoxic edema later. Distal and contralateral edema is attributed to osmotic effects caused by the diffusion of edematous fluid and the accumulation of permeable substances within the bleeding area. Vasogenic cerebral edema results from blood–brain barrier impairment and increased permeability, leading to the leakage of plasma components, including sodium (Na^+) and potassium (K^+) ions. Cytotoxic edema arises from cytotoxic substances disrupting cell energy metabolism, leading to abnormal extracellular ion concentration gradients. Increased extracellular potassium ions are primarily removed through the blood–brain barrier via Na^+-K^+-ATPase-mediated Na^+-K^+ exchange, resulting in a net increase of cations. Our study demonstrated a correlation between lower serum sodium levels and increased edema volume. Potential explanations include the cytotoxic edema perspective: Despite constant serum sodium levels, cytotoxic substances may increase Na^+ in edema fluid while decreasing it in plasma. Lower serum sodium levels may indicate more potent cytotoxic substances, leading to higher edema fluid osmotic pressure and increased edema volume.

Our importance analysis further identified high HDL values, hypertension history, alcohol history, and sex (male) as additional predictors of 72-h brain edema growth risk. Several studies and statistical analyses (34, 35) have reported that patients with cerebral hemorrhage accompanied by poorly controlled hypertension, alcohol consumption, or hyperlipidemia have a significantly higher likelihood of developing severe cerebral edema than healthy individuals. Histological studies have

TABLE 4 Model performance evaluation using training and validation cohorts.

Cohort	Model	AUROC	AUPRC	F1	Sensitivity	Specificity	Accuracy
Training (Mean ± SD)	GDBT	0.603 ± 0.100	0.501 ± 0.112	0.343 ± 0.146	0.839 ± 0.081	0.463 ± 0.187	0.663 ± 0.066
	LR	0.743 ± 0.094	0.630 ± 0.124	0.584 ± 0.106	0.737 ± 0.096	0.546 ± 0.112	0.708 ± 0.076
	RF	0.566 ± 0.100	0.450 ± 0.098	0.354 ± 0.126	0.733 ± 0.100	0.382 ± 0.128	0.609 ± 0.073
	XGBoost	0.575 ± 0.096	0.444 ± 0.091	0.323 ± 0.128	0.779 ± 0.087	0.388 ± 0.150	0.623 ± 0.069
Validation [median (95% CI)]	GDBT	0.654 (0.611, 0.699)	0.548 (0.503, 0.592)	0.444 (0.389, 0.502)	0.769 (0.733, 0.807)	0.526 (0.474, 0.578)	0.615 (0.584, 0.645)
	LR	0.578 (0.535, 0.623)	0.466 (0.423, 0.506)	0.424 (0.374, 0.477)	0.617 (0.573, 0.660)	0.424 (0.376, 0.465)	0.540 (0.497, 0.573)
	RF	0.624 (0.588, 0.687)	0.526 (0.485 , 0.584)	0.413 (0.365, 0.464)	0.667 (0.625, 0.735)	0.438 (0.400, 0.508)	0.556 (0.529, 0.605)
	XGBoost	0.660 (0.611, 0.711)	0.558 (0.501, 0.643)	0.450 (0.391, 0.515)	0.770 (0.720, 0.830)	0.531 (0.456, 0.630)	0.618 (0.575, 0.672)

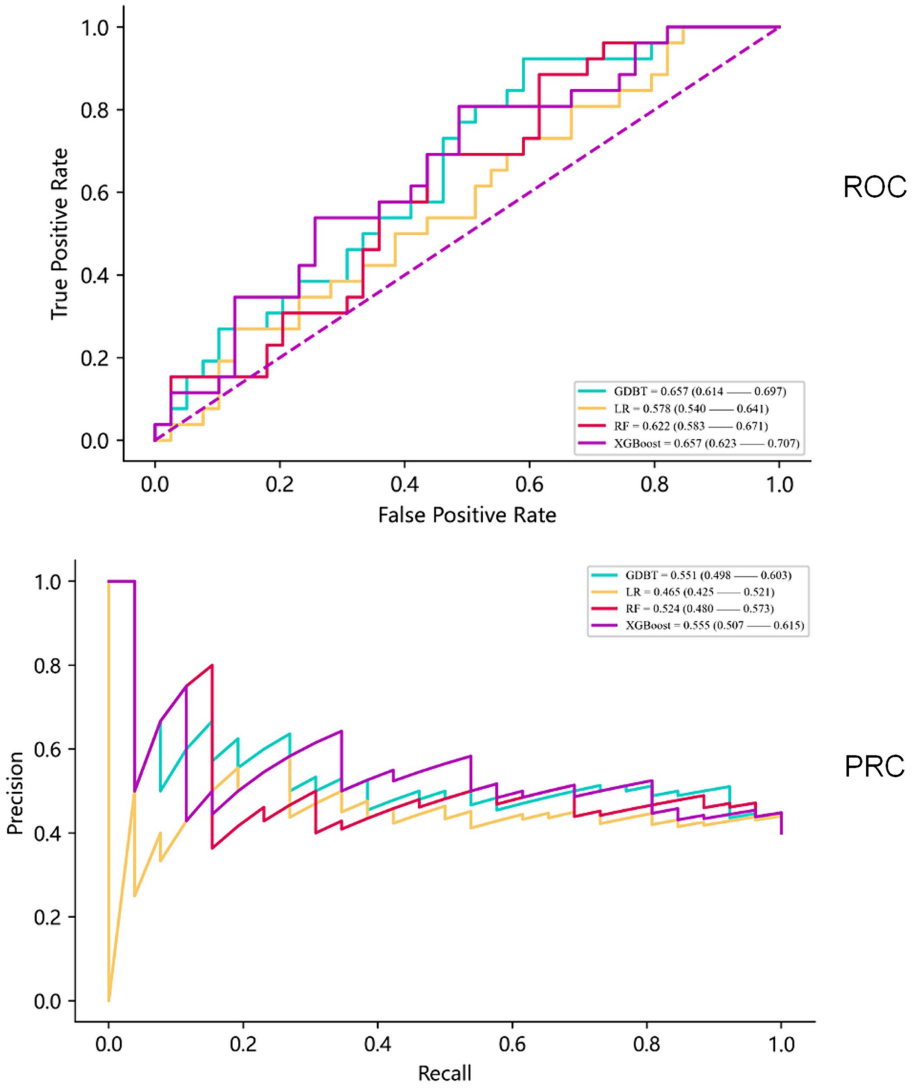


FIGURE 5
Receiver operating curves and precision–recall curves of the models.

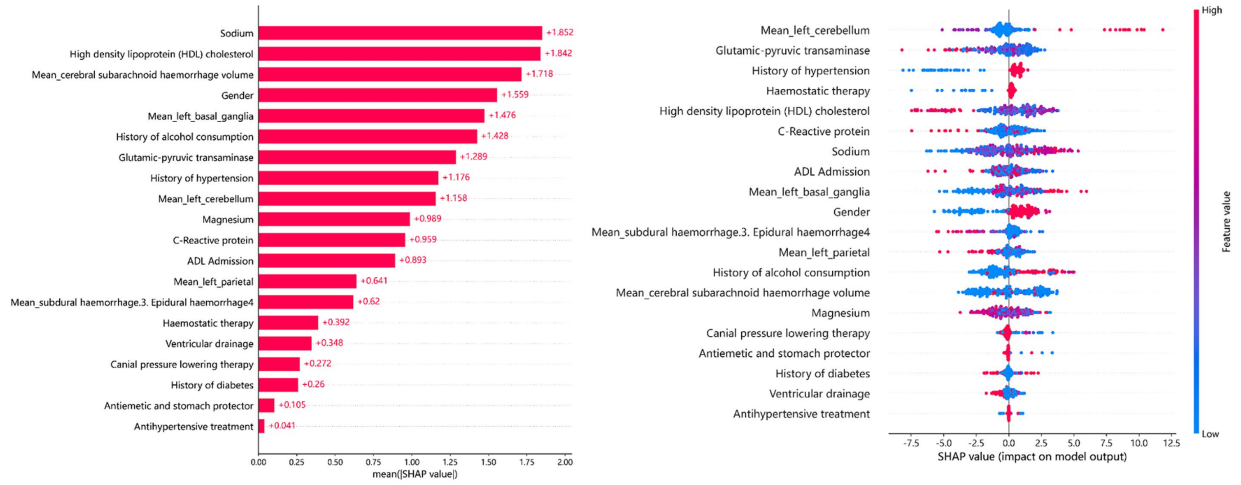


FIGURE 6
SHAP summary plot.

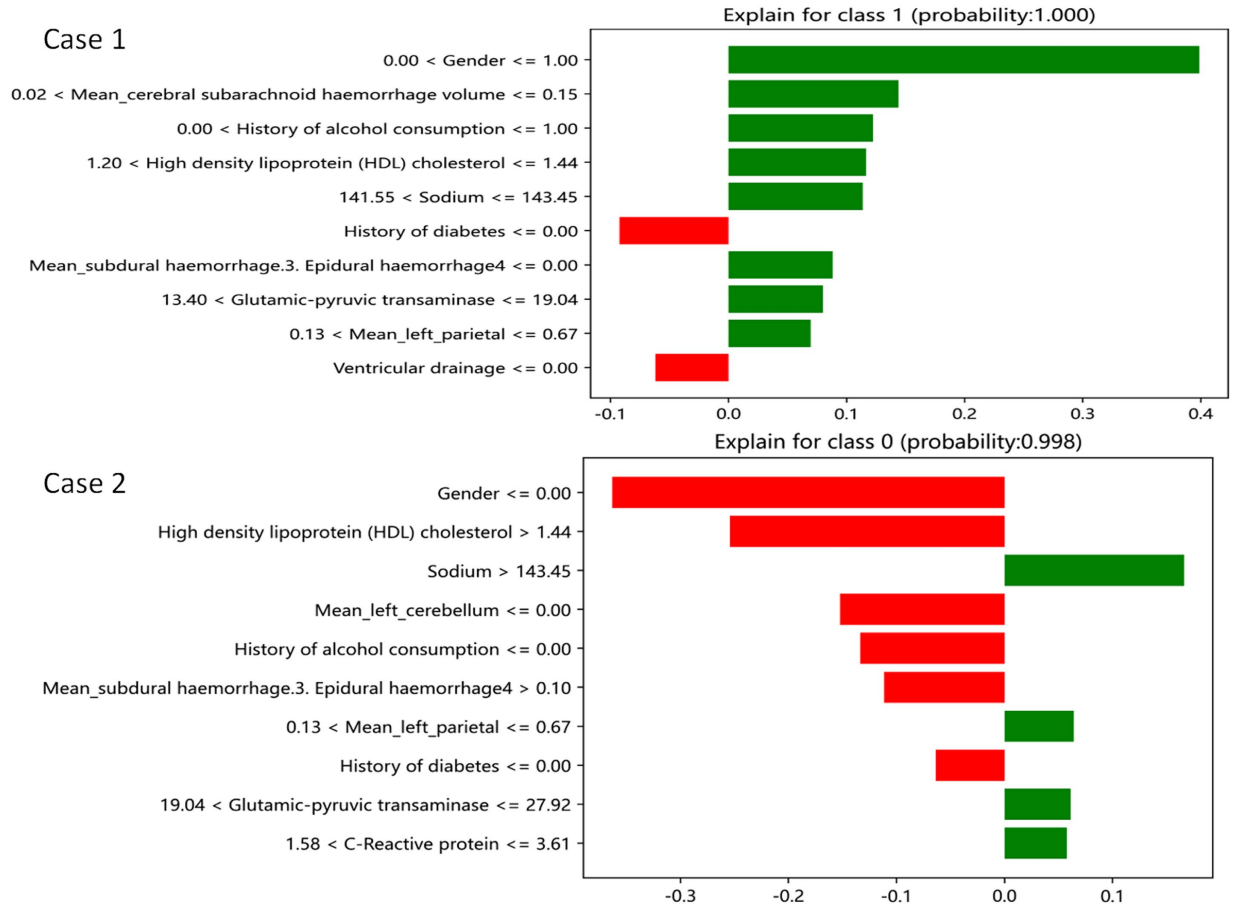


FIGURE 7
GDBT is explained by a locally interpretable model. Features with green bars favor the results, while those with red bars contradict the results. The X axle shows how much of each feature is added or subtracted from the patient's final probability value (i.e., a feature with a weight of 0.3 is equivalent to a 30% change in the probability of the outcome). Class 1 represents increased edema, and class 0 represents no increased edema.

revealed that long-term hypertension can damage small-vessel wall structures. In addition, alcohol consumption, hyperlipidemia, and sex (male) are factors that can exacerbate this damage, contributing to morphological changes associated with cerebral hemorrhage and edema. Therefore, individuals with a history of hyperlipidemia, hypertension, alcohol consumption, or those who are male should be closely monitored for edema growth and receive timely treatment.

Our importance analysis further revealed that the volume of cerebral hematoma was the third most important factor associated with increased edema. In addition, cerebral ventricular drainage, hemostatic treatment, decompressive craniectomy, and antihypertensive treatment effectively reduced edema growth. The hematoma volume is a well-established marker influencing edema volume (36). Although it ranked third in our analysis, we speculate that this might be due to surgical interventions affecting hematoma volume, potentially altering the correlation between hematoma volume and edema growth. This could lead to a less consistent relationship between the two, making hematoma volume less consistently predictive of edema growth. However, further research is needed to confirm this hypothesis. Moreover, the demonstrated effectiveness of ventricle drainage, hemostasis, cranial pressure reduction, and antihypertensive treatment validates the reliability of our predictive model.

Our importance analysis identified CRP and ALT as factors promoting edema growth. Intracerebral hemorrhage is a common clinical condition characterized by rapid onset and progression, posing a significant threat to patient survival. Even if patients survive, they may experience adverse effects on multiple organ functions, leading to multiple organ failure syndrome and death. The primary cause is intracranial hypertension resulting from cerebral hemorrhage, leading to altered consciousness and systemic stress responses. This stimulates various humoral regulatory mechanisms, resulting in strong reactions. CRP and ALT can reflect the severity of the disease, and their elevation suggests a likely deterioration of the patient's underlying condition, increasing the risk of worsening brain edema.

Based on these predictors, the GDBT model developed in this study demonstrated robust and consistent identification and calibration across the training, internal, and external validation cohorts. The selected results were interpretable and could be effectively applied in clinical practice. This model can potentially assist clinicians in identifying high-risk patients and informing clinical decision-making.

Conclusion

The GDBT model consistently demonstrated superior performance in predicting 72-h changes in cerebral edema across the training, internal, and external validation cohorts. The SHAP and LIME analysis revealed that the first three favorable factors associated with increased edema (100%) included the following: sex (male), cerebral subarachnoid hemorrhage volume within the range of 0.02–0.15 mL, and a history of alcohol use. Conversely, the first three favorable factors associated with no increase in edema (99.8%) included the following: serum sodium levels exceeding 143.45 mmol/L, left parietal bleeding volume within the 0.13–0.67 mL range, and alanine aminotransferase levels between 19.04 U/L and 27.92 U/L. These findings have the potential to assist clinicians in the early identification of patients at risk for severe cerebral edema, enabling the implementation of targeted preventive measures to reduce its prevalence.

Strengths

Our study has several advantages, including the inclusion of variables that closely reflect real-world human physiological conditions. The value of the volume of cerebral hemorrhage is collected until the hematoma disappears, no new hematoma occurs during the follow-up period, and the modeling process is a rigorous model development and a validation process, using multiple machine learning algorithms to identify the model with the best performance. Various evaluation measures and model interpretability techniques, such as SHAP and LIME, are used to ensure transparency and facilitate the interpretation of the results, and the model results can be effectively interpreted.

Limitations

Despite its strengths, our study has several limitations. First, the sample size was not validated across multiple centers, and the ROC curve AUC of the constructed model was only 66%, potentially due to the inclusion of many complex and variable factors. The retrospective design of the study and the absence of some data may lead to the exclusion of potentially relevant predictors, such as hypoperfusion due to hypotension, high intracranial pressure, and ischemia or hypoxia due to blood pressure management based on arterial stenosis.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Quzhou Affiliated Hospital of Wenzhou Medical University, Quzhou People's Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JX: Writing – original draft, Writing – review & editing, Conceptualization, Data curation. CY: Conceptualization, Writing – original draft. GY: Data curation, Writing – original draft. HL: Investigation, Writing – review & editing. QD: Conceptualization, Writing – review & editing. DM: Conceptualization, Writing – original draft. CZ: Conceptualization, Writing – original draft. XY: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was funded by Science and Technology Program of QuZhou, China (grant no.2021Y003).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1419608/full#supplementary-material

References

1. Yilmaz C, Kabatas S, Gulsen S, Cansever T, Gurkanlar D, Caner H, et al. Spontaneous supratentorial intracerebral hemorrhage: does surgery benefit comatose patients? *Ann Indian Acad Neurol*. (2010) 13:184–7. doi: 10.4103/0972-2327.70881

2. Zeiser V, Khalaveh F, Cho A, Reinprecht A, Herta J, Rössler K, et al. Risk factors for unfavorable outcome after spontaneous intracerebral hemorrhage in elderly patients. *Clin Neurol Neurosurg*. (2024) 240:108253. doi: 10.1016/j.clineuro.2024.108253

3. Al-Khaled M, Awwad S, Brüning T. Nontraumatic spontaneous intracerebral hemorrhage: baseline characteristics and early outcomes. *Brain Behav*. (2020) 10:e01512. doi: 10.1002/brb3.1512

4. Zaryczańska K, Pawlukowska W, Nowacki P, Zwarzany Ł, Bagińska E, Kot M, et al. Statins and 90-day functional performance and survival in patients with spontaneous intracerebral hemorrhage. *J Clin Med*. (2023) 12:6608. doi: 10.3390/jcm12206608

5. Yao Z, Ma L, You C, He M. Decompressive Craniectomy for spontaneous intracerebral hemorrhage: a systematic review and Meta-analysis. *World Neurosurg*. (2018) 110:121–8. doi: 10.1016/j.wneu.2017.10.167

6. Peng W-J, Li Q, Tang J-H, Reis C, Araujo C, Feng R, et al. The risk factors and prognosis of delayed perihematomal edema in patients with spontaneous intracerebral hemorrhage. *CNS Neurosci Ther*. (2019) 25:1189–94. doi: 10.1111/cns.13219

7. Gu Y, Zhou C, Piao Z, Yuan H, Jiang H, Wei H, et al. Cerebral edema after ischemic stroke: pathophysiology and underlying mechanisms. *Front Neurosci*. (2022) 16:988283. doi: 10.3389/fnins.2022.988283

8. Muscari A, Faccioli L, Lega MV, Lorusso A, Trossello MP, Puddu GM, et al. Predicting cerebral edema in ischemic stroke patients. *Neurol Sci*. (2019) 40:745–52. doi: 10.1007/s10072-019-3717-y

9. Foroushani HM, Hamzehloo A, Kumar A, Chen Y, Heitsch L, Slowik A, et al. Quantitative serial CT imaging-derived features improve prediction of malignant cerebral edema after ischemic stroke. *Neurocrit Care*. (2020) 33:785–92. doi: 10.1007/s12028-020-01056-5

10. He W, Fu X, Chen S. Advancing polytrauma care: developing and validating machine learning models for early mortality prediction. J Transl Med. *J Transl Med*. (2023) 21:664. doi: 10.1186/s12967-023-04487-8

11. Jayatilake S, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. *J Healthc Eng*. (2021) 2021:1–20. doi: 10.1155/2021/6679512

12. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: making sense of the diferences. *Knee Surg Sports Traumatol Arthrosc*. (2022) 30:753–7. doi: 10.1007/s00167-022-06896-6

13. Giordano C, Brennan M, Mohamed B, Rashidi P, Modave F, Tighe P. Accessing artificial intelligence for clinical decision-making. *Front Digit Health*. (2021) 3:645232. doi: 10.3389/fdgth.2021.645232

14. Ngiam KY, Khor IW. Big data and machine learning algorithms for health- care delivery. *Lancet Oncol*. (2019) 20:e262–73. doi: 10.1016/S1470-2045(19)30149-4

15. Ma C, Zhang Y, Niyazi T, Wei J, Guocai G, Liu J, et al. Radiomics for predicting hematoma expansion in patients with hypertensive intraparenchymal hematomas. *Eur J Radiol*. (2019) 115:10. doi: 10.1016/j.ejrad.2019.04.001

16. Wei L, Lin C, Zhou Z, Zhang J, Tan Q, Zhang Y, et al. Analysis of different hematoma expansion shapes caused by different risk factors in patients with hypertensive intracerebral hemorrhage. *Clin Neurol Neurosur*. (2020) 194:105820. doi: 10.1016/j.clineuro.2020.105820

17. Guo DC, Gu J, He J, Chu HR, Dong N, Zheng YF. External validation study on the value of deep learning algorithm for the prediction of hematoma expansion from noncontrast CT scans. *BMC Med Imaging*. (2022) 22:45. doi: 10.1186/s12880-022-00772-y

18. Cao L, Liu M, Wang M, Ding J, Mao K, Liu K, et al. 3D slicer-based calculation of hematoma irregularity index for predicting hematoma expansion in intracerebral hemorrhage. *BMC Neurol*. (2022) 22:452. doi: 10.1186/s12883-022-02983-w

19. Lv CX, An SY, Qiao BJ, Wu W. Time series analysis of hemorrhagic fever with renal syndrome in mainland China by using an XGBoost forecasting model. *BMC Infect Dis*. (2021) 21:839. doi: 10.1186/s12879-021-06503-y

20. Wang J, Xu Y, Liu L, Wu W, Shen C, Huang H, et al. Comparison of LASSO and random forest models for predicting the risk of premature coronary artery disease. *BMC Med Inform Decis Mak*. (2023) 23:297. doi: 10.1186/s12911-023-02407-w

21. Wen Y, Wang X, Li D, Zhang Q, Deng B, Chen Y. Rapid detection of phenytoin sodium by partial-least squares and linear regression models combined with surface-enhanced Raman spectroscopy. *J Pharmaceut Biomed*. (2022) 223:115160. doi: 10.1016/j.jpba.2022.115160

22. Zhou Z, Zare RN. Personal information from latent fingerprints using desorption electrospray ionization Mass spectrometry and machine learning. *Anal Chem*. (2017) 89:1369–72. doi: 10.1021/acs.analchem.6b04498

23. Van den Broeck G, Lykov A, Schleich M, Suciu D. On the tractability of SHAP explanations. *Proc Conf AAAI Artif Intell*. (2021) 35:6505–13. doi: 10.1609/aaai.v35i7.16806

24. Zafar M, Khan N. Deterministic local interpretable model-agnostic explanations for stable. *Explainability Mach learn Knowl Extr*. (2021) 3:525–41. doi: 10.3390/make3030027

25. Hu P, Li Y, Liu Y, Guo G, Gao X, Su Z, et al. Comparison of conventional logistic regression and machine learning methods for predicting delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage: a multicentric observational cohort study. *Front Aging Neurosci*. (2022) 14:857521. doi: 10.3389/fnagi.2022.857521

26. Hines K, Mouchtouris N, al Saiegh F, Hafazalla K, Mazza J, Phan P, et al. Prediction of hematoma expansion in spontaneous intracerebral hemorrhage: our institutional experience. *J Clin Neurosci*. (2021) 86:271–5. doi: 10.1016/j.jocn.2021.01.046

27. Zhang CY, Huang SR, Wang SY, Shen L, Deng BQ. Clinical study of intracranial and extracranial atherosclerotic stenosis in spontaneous intracerebral hemorrhage patients. *J Stroke Cereb Dis*. (2018) 27:286–90. doi: 10.1016/j.jstrokecerebrovasdis.2017.08.011

28. Liu Q, Li X, Wang N, Yang J, Wang K, Chen S, et al. A novel coagulation classification and postoperative bleeding in severe spontaneous intracerebral hemorrhage patients on antiplatelet therapy. *Front Aging Neurosci*. (2022) 14:793129. doi: 10.3389/fnagi.2022.793129

29. Guo R, Zhang R, Liu R, Liu Y, Li H, Ma L, et al. Machine learning-based approaches for prediction of Patients' functional outcome and mortality after spontaneous intracerebral hemorrhage. *J Pers Med*. (2022) 12:112. doi: 10.3390/jpm12010112

30. Wen M, Ye J, Han Y, Huang L, Yang H, Jiang W, et al. Hypertonic saline regulates microglial M2 polarization via mi R-200b/KLF4 in cerebral edema treatment. *Biochem Biophys Res Commun*. (2018) 499:345–53. doi: 10.1016/j.bbrc.2018.03.161

31. Chen S-Y, Thompson-Leduc P, Sawyer RN, Fakih I, Cheung HC, Macheca M, et al. Outcomes and resource use of patients with large hemispheric infarction and cerebral edema: analysis of real-world data. *Curr Med Res Opin*. (2021) 37:781–8. doi: 10.1080/03007995.2021.1900090

32. Doron O, Zadka Y, Barnea O, Rosenthal G. Interactions of brain, blood, and CSF: a novel mathematical model of cerebral edema. *Fluids Barriers CNS*. (2021) 18:42. doi: 10.1186/s12987-021-00274-z

33. Hsia DS, Tarai SG, Alimi A, Coss-Bu JA, Haymond MW. Fluid management in pediatric patients with DKA and rates of suspected clinical cerebral edema. *Pediatr Diabetes*. (2015) 16:338–44. doi: 10.1111/pedi.12268

34. Miao J, Song X, Sun W, Qiu X, Lan Y, Zhu Z. Predictors of malignant cerebral edema in cerebral artery infarction: a meta-analysis. *J Neurol Sci*. (2019) 409:116607. doi: 10.1016/j.jns.2019.116607

35. Wang Y-F, Parpura V. Astroglial modulation of Hydromineral balance and cerebral edema. *Front Mol Neurosci*. (2018) 11:204. doi: 10.3389/fnmol.2018.00204

36. Kang J, Huang Q, Liu Y. Advance in research on the genetic etiology of spontaneous intracerebral hemorrhage. *Chin J Med Genet*. (2016) 33:702–7. doi: 10.3760/cma.j.issn.1003-9406.2016.05.028

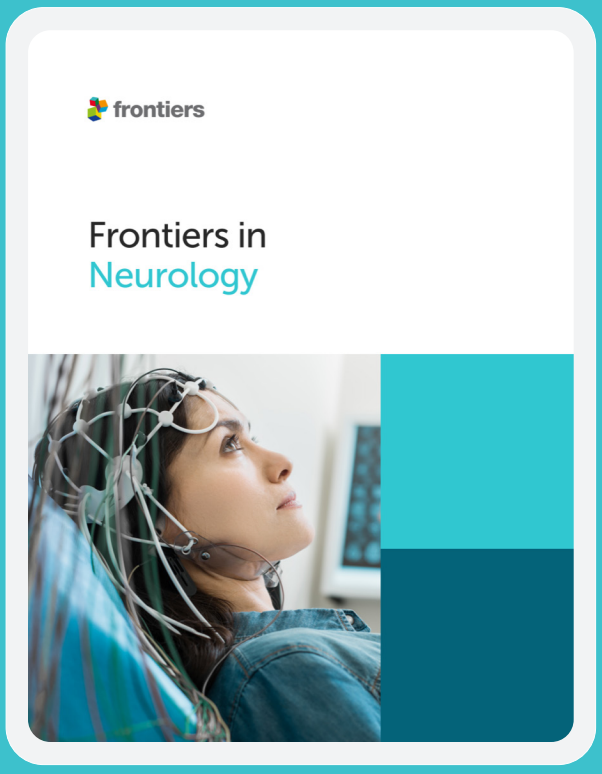
Frontiers in Neurology

Explores neurological illness to improve patient care

The third most-cited clinical neurology journal explores the diagnosis, causes, treatment, and public health aspects of neurological illnesses. Its ultimate aim is to inform improvements in patient care.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact