frontiers | Research Topics

The application of artificial intelligence in interventional neuroradiology

Edited by
Yuhua Jiang, Zhongrong Miao, Youxiang Li, Ferdinand Hui and Baofeng Gao

Published in
Frontiers in Neurology

ISSN 1664-8714
ISBN 978-2-8325-2859-4
DOI 10.3389/978-2-8325-2859-4

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

The application of artificial intelligence in interventional neuroradiology

Topic editors

Yuhua Jiang — Capital Medical University, China
Zhongrong Miao — Capital Medical University, China
Youxiang Li — Capital Medical University, China
Ferdinand Hui — The Johns Hopkins Hospital, Johns Hopkins Medicine, United States
Baofeng Gao — Beijing Institute of Technology, China

Citation

Jiang, Y., Miao, Z., Li, Y., Hui, F., Gao, B., eds. (2023). *The application of artificial intelligence in interventional neuroradiology*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2859-4

Table of contents

frontiers | Frontiers in Neurology

TYPE Editorial
PUBLISHED 05 January 2023
DOI 10.3389/fneur.2022.1112624

OPEN ACCESS

EDITED AND REVIEWED BY
Ulises Gomez-Pinedo,
Health Research Institute of Hospital Clínico San Carlos, Spain

*CORRESPONDENCE
Yuhua Jiang
jxy_200321@163.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Experimental Therapeutics, a section of the journal Frontiers in Neurology

RECEIVED 30 November 2022
ACCEPTED 16 December 2022
PUBLISHED 05 January 2023

CITATION
Jiang Y, Lv J, Li Y and Zhang Y (2023) Editorial: The application of artificial intelligence in interventional neuroradiology.
Front. Neurol. 13:1112624.
doi: 10.3389/fneur.2022.1112624

Editorial: The application of artificial intelligence in interventional neuroradiology

Yuhua Jiang[1,2*†], Jian Lv[1,2†], Youxiang Li[1,2] and Yanling Zhang[3]

[1]Department of Neurosurgery, Beijing Neurosurgical Institute and Beijing Tiantan Hospital, Capital Medical University, Beijing, China, [2]Department of Neurointerventional Engineering and Technology, Beijing Engineering Research Center, Beijing, China, [3]Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

KEYWORDS

interventional neuroradiology, artificial intelligence, cerebrovascular disease, cerebrovascular interventional therapy, prediction models

Editorial on the Research Topic
The application of artificial intelligence in interventional neuroradiology

Introduction

As editors of this Research Topic, it was our pleasure to introduce novel findings and new achievements in the application of artificial intelligence (AI) in interventional neuroradiology. Cerebrovascular disease is becoming an increasingly important cause of death and intervention therapy has become the mainstay treatment for this disease. However, there is a need to assess the efficacy and safety of endovascular therapy. In recent years, AI technology has advanced rapidly and has shown great promise in solving complicated problems. It also possesses strong potential to improve the clinical application of cerebrovascular interventional therapy.

This Research Topic consists of 10 papers, including seven original research articles, one system review, one mini review, and one commentary. The purpose of this Editorial is to summarize the key findings and perspectives presented in each of the accepted articles.

The basic work of applying the AI technology in practice is to design and establish a database. So, You et al. introduced a protocol for constructing a multicenter database based on CTA images of IAs. This protocol described how to collect research data, conduct aneurysm segmentation, and annotation. This study exemplifies the paradigm of building a big database for artificial intelligence.

A similar study has been performed previously (1). From such studies we know that a well-established database of the AI model can directly improve the performance of the model.

Compared to statistical regression models, machine learning (ML) models, which are part of AI technology, can identify non-intuitive patterns in variables which might be missed by statistical tests. Endovascular treatment strategies can be optimized by employing ML models to predict therapeutic outcomes.

However, clinicians lack the skills to handle ML data scientific, which hinder the development of such models. Ou et al. designed three models for predicting IAs endovascular treatment outcome. They created an automated machine learning (AutoML)-derived model which showed better performance compared with the manually trained ML and statistical models. The AutoML model has the advantage of simplifying the process of building a ML model without relying on experts. Therefore, it can be used by people without expertise in artificial intelligence. However, these studies had several limitations. For instance, they did not assess the calibration of various algorithms as mentioned in the commentary by Huber et al.

Assessment of clinical outcomes is essential for acute ischemic stroke (AIS) patients. Jabal et al. built ML models for pre-intervention prediction of the 90-day dichotomized modified Rankin Scale (mRS-90) scores for AIS patients who underwent thrombectomy, using clinical and radiological information extracted from CTA and CTA with the e-stroke software. The authors used various ML algorithms including *k*-nearest neighbors, random forests (RF), gradient boosting (GB), and Extreme Gradient Boosting (XGBoost) and found that XGBoost was the best performing classifier.

Although endovascular treatment has become the mainstay treatment for IAs, its related complications should not be ignored. Clinicians should balance between the risk of complications from endovascular treatment and the risk of IAs rupture. Tian et al. constructed several ML prediction models to study endovascular procedure-related complications of IAs and found that the ANN models had the best performance. In the study, 443 patients were enrolled, and the three most significant features were distal aneurysm, aneurysm size, and treatment modality as determined by the Shapely Additive explain (SHAP) and feature importance analysis.

Pipeline embolization device (PED) is the most commonly used flow diverter service for the treatment of IAs. In-stent stenosis (ISS) is a common complication after PED placement and might adversely affect long-term prognosis. Wei et al. built ML prediction models using clinical, laboratory, and imaging data obtained from 435 patients. They compared the prediction performance of five ML algorithms including elastic net (ENT), SVM, XGBoost, Gaussian Naïve bayes, and random forest. Through SHAP analysis, they found that internal carotid artery location was the most important predictor.

Notably, AI prediction models cannot override traditional statistical models totally because of the data and technology limitations. The continuous advancement of AI and medical image processing technology, is expected to provide multi-dimensional information upon which precise AI prediction models will be built. Similar opinions were held by Zeng et al. In their systematic review, they enrolled 16 articles, including 19 ML and DL models for predicting prognosis of stroke patients with large vessel occlusion (LVO). They found that AI did not show an overall advantage over existing prognostic scores. Therefore, whether ML and DL methods can improve prediction of stroke outcomes in LVO still need to be further clarified. Marasini et al. reviewed AI methods applied in IAs detection, IAs screening, IAs rupture prediction, IAs clinical decision support and workflow enhancement of IAs. They reported that AI can handle large numbers of variables and identify non-linear relationships among them. However, despite significant advances in the field of AI, its application in many areas, particularly in non-imaging data is at the foundational stage. Therefore, further studies are needed to improve the prediction performance of AI for periprocedural complications related to endovascular therapy.

Author contributions

YJ and JL wrote the original manuscript. YL and YZ reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We would like to thank all authors of this Research Topic for their valuable contributions. We also thank the Editorial board of Frontiers in Neurology for their support.

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

References

1. Shi Z, Miao C, Schoepf UJ, Savage RH, Dargis DM, Pan C, et al. A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. *Nat Commun*. (2020) 11:6090. doi: 10.1038/s41467-020-19527-w

frontiers in Neurology

ORIGINAL RESEARCH
published: 29 November 2021
doi: 10.3389/fneur.2021.735142

Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study

Chubin Ou[1,2†], Jiahui Liu[1†], Yi Qian[1], Winston Chong[3], Dangqi Liu[4], Xuying He[1], Xin Zhang[1] and Chuan-Zhi Duan[1*]*

[1] Neurosurgery Center, Department of Cerebrovascular Surgery, The National Key Clinical Specialty, The Engineering Technology Research Center of Education Ministry of China on Diagnosis and Treatment of Cerebrovascular Disease, Guangdong Provincial Key Laboratory on Brain Function Repair and Regeneration, The Neurosurgery Institute of Guangdong Province, Zhujiang Hospital, Southern Medical University, Guangzhou, China, [2] Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW, Australia, [3] Monash Medical Centre, Monash University, Clayton, VIC, Australia, [4] Department of Neurosurgery, The First People's Hospital of Foshan, Foshan, China

OPEN ACCESS

Edited by:
Yuhua Jiang,
Capital Medical University, China

Reviewed by:
Luis Rafael Moscote-Salazar,
Latinamerican Council of Neurocritical Care (CLaNi), Colombia
Jianwei Pan,
Zhejiang University, China

***Correspondence:**
Xin Zhang
zhangxin19830818@163.com
Chuan-Zhi Duan
doctor_duanzj@163.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

Received: 02 July 2021
Accepted: 01 November 2021
Published: 29 November 2021

Citation:
Ou C, Liu J, Qian Y, Chong W, Liu D, He X, Zhang X and Duan C-Z (2021) Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study. Front. Neurol. 12:735142. doi: 10.3389/fneur.2021.735142

Background: The prediction of aneurysm treatment outcomes can help to optimize the treatment strategies. Machine learning (ML) has shown positive results in many clinical areas. However, the development of such models requires expertise in ML, which is not an easy task for surgeons.

Objectives: The recently emerged automated machine learning (AutoML) has shown promise in making ML more accessible to non-computer experts. We aimed to evaluate the feasibility of applying AutoML to develop the ML models for treatment outcome prediction.

Methods: The patients with aneurysms treated by endovascular treatment were prospectively recruited from 2016 to 2020. Treatment was considered successful if angiographic complete occlusion was achieved at follow-up. A statistical prediction model was developed using multivariate logistic regression. In addition, two ML models were developed. One was developed manually and the other was developed by AutoML. Three models were compared based on their area under the precision-recall curve (AUPRC) and area under the receiver operating characteristic curve (AUROC).

Results: The aneurysm size, stent-assisted coiling (SAC), and posterior circulation were the three significant and independent variables associated with treatment outcome. The statistical model showed an AUPRC of 0.432 and AUROC of 0.745. The conventional manually trained ML model showed an improved AUPRC of 0.545 and AUROC of 0.781. The AutoML derived ML model showed the best performance with AUPRC of 0.632 and AUROC of 0.832, significantly better than the other two models.

Conclusions: This study demonstrated the feasibility of using AutoML to develop a high-quality ML model, which may outperform the statistical model and manually derived ML models. AutoML could be a useful tool that makes ML more accessible to the clinical researchers.

Keywords: intracranial aneurysm, AutoML, endovascular treatment, machine learning, stroke

INTRODUCTION

Endovascular therapy is widely used in the treatment of intracranial aneurysms (1). Despite a remarkable advancement of the endovascular coiling for intracranial aneurysms, there still exists a high rate of recurrence and recanalization. It has been reported that the recanalization rate for coiling and flow diversion are 20.8 and 10.2%, respectively (2, 3). Approximately up to 50% of patients who succumbed to recurrence or recanalization necessitated further treatment, which may place an additional financial burden on the patients. Moreover, recanalization puts patients at increased risk of a thromboembolic event or aneurysm rupture. Many studies have tried to study the risk factors for recanalization. The aneurysm size, morphologies, treatment strategies, and hemodynamics have been found to be associated recanalization (4–9). Some studies have tried to develop the models or grading scales to predict treatment outcome (4, 10–12). However, evaluation of some of the grading scales showed relatively poor performance (13).

In recent years, the machine learning (ML) models, as an alternative to the conventional statistical model, have shown promise in many clinical areas (10, 14, 15). ML models can learn complex relationships from a large amount of data. Compared with a regression model that focus on statistically significant variables, the ML models can discover non-intuitive patterns from variables which may be overlooked by statistical test (16).

Although the ML models have shown outstanding performance, the development of such models requires expertise in ML. Despite the existence of open-source code libraries, such as Scikit-Learn, PyTorch, and Tensorflow, their use still requires significant experience in programming and knowledge of ML. In addition, a high-quality model usually requires expertise to tune and train. All these problems pose a great challenge for the clinical researchers hoping to adopt ML in their research.

The recently emerged automated machine learning (AutoML) has found a way to close the gap between ML and non-artificial intelligence (non-AI) experts. The emergence of AutoML automates the process of building an ML model which in the past relied on data scientists. This lowers the learning threshold for using ML and allows people without expertise in ML to apply ML to their own area. It has recently been reported that AutoML has helped the physicians to develop the ML models that achieved good performance in the field of medical image analysis and disease risk prediction (17, 18). However, such success has not been reported in the field of stroke treatment.

Therefore, in this study, we aimed to evaluate the feasibility of using AutoML to develop the ML models for aneurysm treatment outcome prediction. Treatment was considered successful if angiographic complete occlusion was achieved at follow-up. We developed the prediction models for treatment outcome using three different methods: a statistical multivariate regression model, a manually derived ML model, and an AutoML derived ML model, and compared their performance.

Abbreviations: AutoML, automatic machine learning; ML, machine learning; AUPRC, area under precision-recall curve; AUROC, area under receiver-operating characteristic curve; SAC, stent-assisted coiling; FD, flow diversion.

METHODS

Patient Cohorts

The patients were recruited according to the protocol of a prospective cohort (19). The primary endpoints of the cohort study are an evaluation of the safety and efficacy of interventional treatment for 6 months after surgery, with each participant completing at least 1 year of follow-up. Approval for this study was obtained from the local Institutional Review Board. The data used in the current study were anonymous and the requirement for informed consent was therefore waived. From the prospective cohort, we included the aneurysm cases treated by endovascular treatment. Dissecting aneurysms and fusiform aneurysms, aneurysms with prior treatment, or the cases with missing clinical information were excluded. A total of 395 patients were identified from our center. However, due to loss to follow-up or incomplete record, only 182 patients and 218 aneurysms with complete record of angiographic follow-up were used in the current study.

Data Acquisition

The morphological parameters were measured and calculated from three-dimensional digital subtraction angiography (DSA) images prior to treatment. The measurements were done by two independent neurosurgeons and the average of their readings were used. The clinical symptoms, such as feeling of headache, nausea, vomit, and dizziness were recorded. The blood tests were also performed for the patients prior to treatment to measure lipid level and blood clotting function. Additionally, the patient demographics, medical history, and lifestyle behaviors were recorded. Treatment related parameters, such as treatment method, number of coils stent metal coverage rage (MCR) were included. Immediate angiographic outcome after treatment and follow-up angiographic outcome were also recorded according to the Raymond–Roy Occlusion Classification scale (20). Treatment was considered successful if complete occlusion was achieved at follow-up. The average follow-up time for the coiling and stent-assisted coiling (SAC) cases is 9.4 and 14.2 months for flow diversion cases. The complete list of collected variables is shown in **Table 1**.

General Procedures of ML

The general procedures of ML include the following steps: feature selection, feature engineering, ML model selection, and hyperparameter tuning, as shown in **Figure 1**. In feature selection, the features that are relevant to the prediction target are selected based on various criteria, such as ANOVA F-value, chi-squared statistics, univariate statistical significance *P* value, and information gain. Feature selection help to identify and focus on the useful features. In feature engineering, raw features can be normalized, binarized, decomposed, or combined to create new features, which might help to better model the data. In model selection, various ML

TABLE 1 | Result of univariate analysis.

Variable	Occluded (*N* = 194)	Recanalized (*N* = 24)	*P*
Gender (Female)	127	15	0.748
Age	54.9 ± 10.9	54.4 ± 11.7	0.931
Dizzy	90	9	0.409
Headache	110	11	0.312
Nausea	159	15	0.025*
Vomit	161	15	0.016*
Alcohol	25	5	0.286
Smoking	26	4	0.661
Labor work	10	1	0.835
Lack of sleep	21	5	0.297
Height	161.7 ± 7.5	164.7 ± 7.4	0.137
Weight	59.7 ± 9.0	58.3 ± 9.6	0.586
Systole	130.6 ± 18.6	129.1 ± 20.1	0.766
Diastole	80.8 ± 10.1	79.3 ± 12.3	0.890
Glucose	5.5 ± 1.8	5.6 ± 1.6	0.455
GHb	5.8 ± 0.7	5.7 ± 0.6	0.571
WBC	7.3 ± 3.1	7.7 ± 3.3	0.378
Platelet	239.0 ± 57.9	261.6 ± 65.9	0.141
Triglyceride	1.4 ± 1.4	1.1 ± 0.6	0.144
Cholesterol	4.6 ± 0.9	4.4 ± 0.8	0.639
LDL	2.7 ± 0.8	2.9 ± 0.9	0.443
HDL	1.3 ± 0.3	1.3 ± 0.3	0.997
Fibrin	3.3 ± 0.8	3.6 ± 1.1	0.238
APTT	35.7 ± 3.5	36.9 ± 3.8	0.169
PT	12.9 ± 0.7	13.1 ± 0.9	0.771
Hcy	10.7 ± 3.9	11.1 ± 6.1	0.531
Multiple	74	10	0.738
Rupture	161	20	0.966
Hypertension	55	9	0.530
ICA	121	11	0.118
MCA	26	2	0.484
ACA and AComA	26	4	0.661
PComA	12	2	0.686
Posterior circulation	9	5	0.002*
Irregular shape	33	8	0.054
Aneurysm size	4.9 ± 3.2	7.8 ± 4.5	0.003*
Sac width	4.6 ± 3.2	7.0 ± 4.5	0.013*
Sac height	4.4 ± 2.9	6.9 ± 4.2	0.009*
Neck width	4.0 ± 1.8	4.6 ± 2.2	0.094
Vessel angle	100.1 ± 29.2	111.2 ± 36.0	0.232
Parent artery	3.1 ± 0.9	3.1 ± 0.8	0.699
Size ratio	1.7 ± 1.3	2.5 ± 1.5	0.006*
Aspect ratio	1.1 ± 0.6	1.5 ± 0.8	0.011*
Previous SAH	20	0	0.256
SAC	137	9	<.001*
FD	19	0	0.109
Neck metal coverage	17.5 ± 10.5	12.8 ± 5.7	0.161
Post-procedure Angiographic Occlusion	30	4	0.023*
mRS	0.62 ± 1.01	0.51 ± 0.77	0.116

**indicates statistical significance, P < 0.05.*

algorithms are evaluated on the dataset and the best is selected. Common ML algorithms, to list a few, include Support Vector Machine, K-Nearest Neighbors, Decision Tree, Artificial Neural Network, Random Forest, and Naïve Bayes. All these algorithms have a wide range of hyperparameters that require careful adjustment to suit different tasks and datasets. For example, Random Forest have more than a dozen of hyperparameters, such as maximum number of tresses, maximum tree depth, maximum number of features, and minimum samples in leaf. In hyperparameters tuning, the optimal hyperparameters are usually found using grid-search or randomized grid-search over millions of possible combinations of hyperparameters.

A pipeline consists of a combination of specific methods for feature selection, feature engineering, ML algorithm, and a specific set of hyperparameters. To build a good ML model, one needs to identify an optimal pipeline that achieves best performance on the dataset.

Automated Machine Learning

Automated machine learning automates the above pipelines and explores different choices of algorithms, feature selection and feature engineering technique, and hyperparameters. Since each major step in the pipeline involves dozens to millions of choices, complete exploration of all possible pipelines is inefficient and impractical. To speed-up the search process, we employed an algorithm called Tree-based Pipeline Optimization (TPOT) to automate the pipeline search. TPOT is based on the evolutionary algorithm which uses genetic programming to search for optimal pipeline (21). Genetic programming mimics the way of natural selection. Briefly, in each optimization run (generation), TPOT randomly generates multiple pipelines (population). These pipelines were evaluated based on their accuracy (fitness to survive). The best few pipelines (scored by accuracy) were selected into the next optimization run (selection). The selected pipelines were then randomly modified (mutation and crossover) in which a few of the pipeline elements (e.g., ML models, feature selection methods, and feature processing method) are changed. Several generations are run and the pipeline that performed best on the training set is selected as the optimal pipeline.

In the current study, AutoML was used on the training set to obtain an optimal pipeline. To avoid overfitting, 10-fold cross-validation was used. For the setting of AutoML, the number of generations to run was set to 10 and the population size at each generation was 100. Increasing the number of generations or the population size can result in higher chance of discovering better pipelines but at the cost of computational time. In the current study, the program was run on a desktop computer (CPU: Intel i7 8700) for ∼1 h.

After obtaining the optimal pipeline, the derived model was evaluated on the test set. To further avoid overoptimistic results due to random split of the training and test set, the above procedures were repeated 20 times and each time with a different split of training and test set. The average performance from the 20 repeats was reported. The training and evaluation procedures are shown in **Figure 2**.

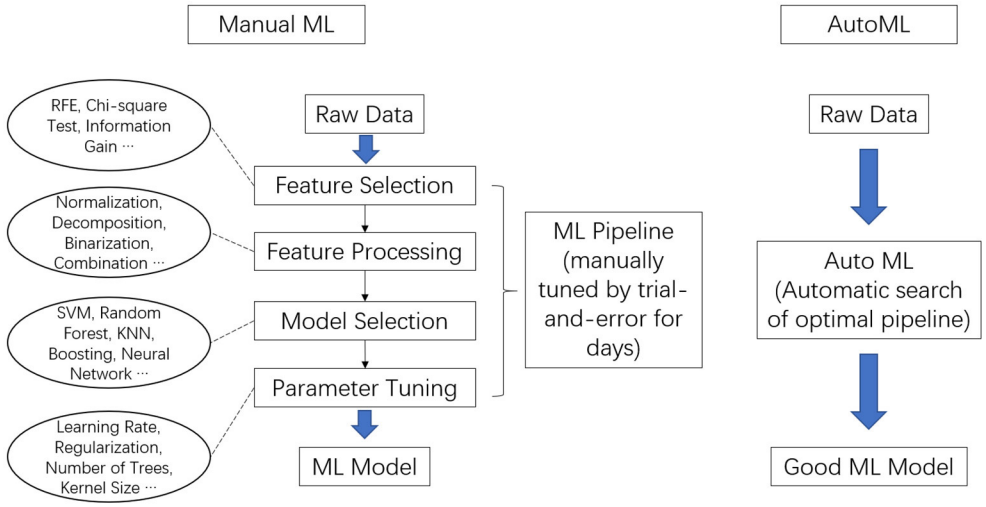


FIGURE 1 | General pipeline of training a machine learning **(ML)** model (left) and training using an automated machine learning (AutoML) (right).

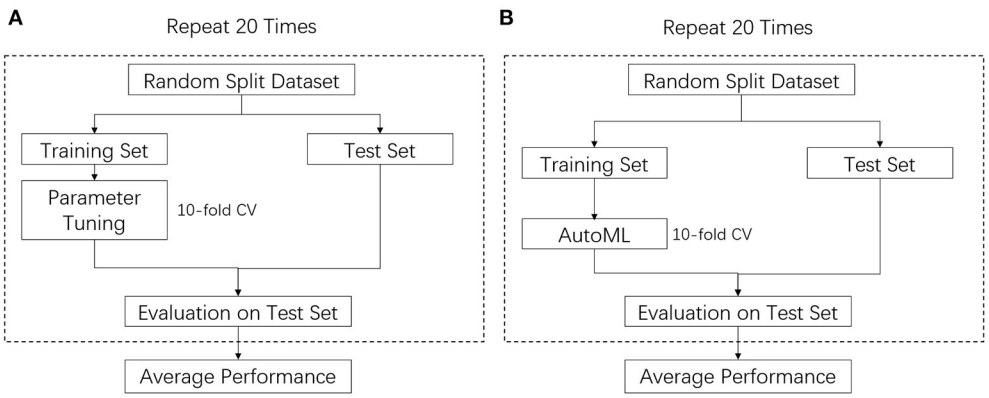


FIGURE 2 | Training and evaluation procedures for manual ML **(A)** and AutoML **(B)**.

Conventional ML

For comparison purposes, an ML model was trained manually using a typical method found in the literature. Random forest is one of the most popular algorithms used in the literature and is often found to have better performance than logistic regression (22–24). To represent a typical scenario in which a non-ML-expert develops an ML model for clinical research, we applied the same training procedures as described in the work of Rubber et al. The manual pipeline started with feature processing using normalization, and model building using the random forest algorithm. The hyperparameter of the algorithm (number of trees) was tuned between 5 and up to 5,000 (24). The hyperparameters were tuned using 10-fold cross-validation on the training set and the model was tested on the test set. The above procedures were repeated 20 times and each time with a different split of training and test set. The average performance of the 20 repeats was reported. The training and evaluation procedures that were used are shown in **Figure 2**.

Statistical Model Building

All variables of the successful and unsuccessful cases were compared using the univariate analyses. For binary or categorical variables, the Fisher's exact test or chi-square test was performed. For continuous variables, they were first examined with the Shapiro–Wilk test to determine normality, followed by Student's *t*-test (for normally distributed variables) or Mann–Whitney *U*-test (for non-normally distributed variables). The variables with $P < 0.05$ in the univariate analysis were further selected into multivariate analysis using a backward conditional stepwise method. The statistical analyses were performed using SPSS (IBM Corporation, NY, USA). The variables that remained statistically significant ($P < 0.05$) in multivariate analysis were used for the statistical model building. For a fair comparison with other methods, a logistic regression model was fitted on the training set and evaluated on the test set. The training and evaluation procedures were also repeated 20 times and each time with a different split of training and test set. The average performance from the 20 repeats was reported.

TABLE 2 | Result of multivariate analysis.

Variable	OR	*P*
Aneurysm size	1.242 (95% CI 1.090-1.416)	0.001
SAC	0.208 (95% CI 0.079-0.546)	0.001
Posterior circulation	4.383 (95% CI 1.046-18.370)	0.043

TABLE 3 | Summary of model performance.

	Statistical	Manual ML	Auto ML	ARSS
Sensitivity	1.000	1.000	1.000	1.000
PPV	0.167	0.342	0.408	0.142
AUROC	0.745	0.781	0.823	0.771
AUPRC	0.432	0.545	0.632	0.496
F1-score	0.286	0.508	0.578	0.378

Aneurysm Recanalization Stratification Scales (ARSS)

For comparison with the currently used method, we chose the Aneurysm Recanalization Stratification Scales (ARSS) proposed by Ogilvy et al. (25). The scale was calculated by assigning different weights to different risk factors. Aneurysm-specific factors include size (>10 mm), 2 points; rupture, 2 points; presence of thrombus, 2 points. Treatment-related factors include stent assistance, −1 point; flow diversion, −2 points; Raymond-Roy 2 occlusion, 1 point; Raymond-Roy 3 occlusion, 2 points. We evaluated the same test set used in the other three methods for assessing the averaged performance in 20 repeats.

Model Comparison

Though unsuccessful cases only consist of a small portion (11%) of the dataset, it is more important to identify the unsuccessful cases than the successful cases. To avoid bias introduced by imbalanced data, besides the commonly used receiver operating characteristic (ROC) curve, we also used the precision-recall curve (area under the precision-recall curve [AUPRC]) as the evaluation metric, which is more informative than ROC when evaluating classifier on imbalanced data (26). The precision-recall curve plots precision, also termed as positive predictive value (PPV), against recall (sensitivity). The AUPRC is a balanced measure of the capability of a model to predict unsuccessful cases. The comparison of the performances of three models in the 20 repeats was examined by Wilcoxon signed ranks test as suggested by a previous study (27).

RESULTS

A total of 182 patients with 218 aneurysms were included. The average aneurysm size was 5.3 mm. The majority of them were located on the internal carotid artery (ICA), followed by the middle cerebral artery (MCA) and anterior communicating artery (AComA). At follow-up, only 24 cases remained unoccluded. The baselines for the successfully treated and unsuccessfully treated group are summarized in **Table 1**. In the univariate analysis, aneurysm size, aneurysm width, aneurysm height, presence of nausea, presence of vomit, use of SAC, aneurysm location in the posterior circulation, and the immediate post-procedure angiographic outcome showed statistical significance. In the multivariate analysis, only aneurysm size, use of SAC, and posterior circulation remained as significant variables, as shown in **Table 2.**

The sensitivity, positive predictive value, area under the receiver operating characteristic curve (AUROC), AUPRC, and F1-score of the three models are summarized in **Table 3**. The statistical model achieved an AUPRC of 0.432 (95% *CI* 0.373–0.491), as shown in **Figure 3**. The manually derived ML model achieved better performance, with a value of 0.545 (95% *CI* 0.458–0.632). The ARSS model achieved an AUPRC of 0.496 (95% *CI* 0.418–0.574). The AutoML derived model achieved the best performance with an AUPRC of 0.632 (95% *CI* 0.585–0.679). The AUPRC of AutoML derived model was significantly higher than that from the statistical model ($P < 0.001$) and that from manual derived ML model ($P = 0.021$) and that from the ARSS model ($P = 0.011$).

The procedures of applying AutoML in clinical settings are shown in **Figure 4**. The surgeons first prepare data and then run the few lines of code of AutoML and get an automatically generated Python file that contains the optimal pipeline to build a high-quality ML model. The surgeons can then use the generated python code to train an ML model and predict the risk of recanalization. In the current study, the optimal pipeline obtained started with feature selection using recursive feature elimination with Extra-Trees classifier, followed by feature preprocessing using Normalization. The algorithm used to build the model was the Gradient-Boosting classifier.

DISCUSSION

Recanalization and recurrence are the Achilles' Heel of endovascular treatment. This can only be confirmed by a long-term follow-up study. Thus, the question is raised: are there any methods to predict the long-term outcome of embolization? Recently, the ML models have emerged as alternatives to the traditional statistical models used to predict disease risk and therapeutic effect. However, ML is often recognized as complicated technology accessible only to a small fraction of medical researchers and data scientists. The advantage of AutoML is that it allows non-ML experts to utilize the ML models without prior expertise. In this study, we found that AutoML, with the only minimum amount of code, could develop an ML model that performed significantly better than the commonly used statistical model in predicting treatment success.

Comparison of the AutoML Model and Statistical Model

While the statistical models are easy to derive and understand, they have several limitations. They assume linear independence between the variables which may fail to account for interactions

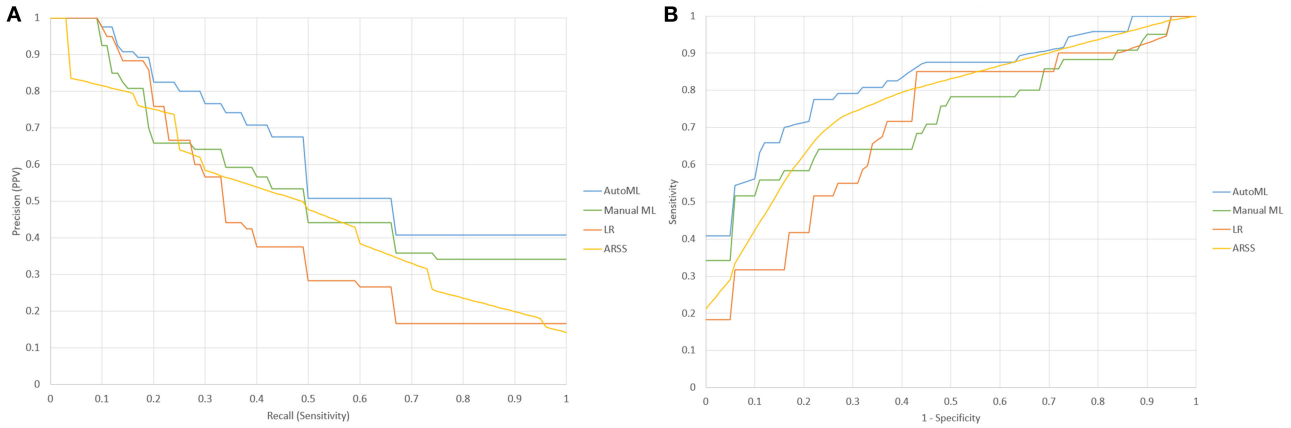


FIGURE 3 | (A) Precision-recall characteristic curves of a statistical model **(LR)**, manually derived ML model (Manuel), AutoML derived ML model (AutoML), and Aneurysm Recanalization Stratification Scale **(ARSS)**; **(B)** the receiver operating characteristic **(ROC)** curves of statistical model **(LR)**, ManualML, AutoML, and ARSS.

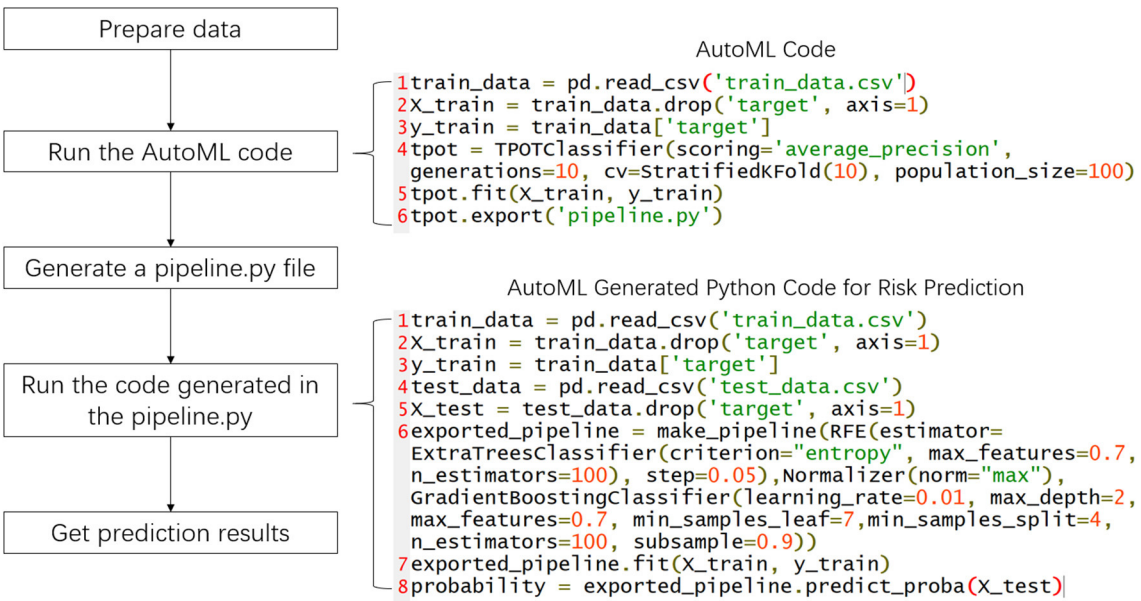


FIGURE 4 | General procedures to apply AutoML in the clinical settings.

between the variables. The prescreening of variables using *P* values may also miss important variables which may not appear statistically significant in a univariate test (28). In contrast, the ML models can learn nonlinear and interactive patterns between variables and thus producing a more accurate prediction model. Many studies have reported that an ML model outperformed the statistical models (22–24). However, there are several drawbacks that limit the use of the ML model in clinical research. One is the black-box problem of an ML algorithm yet this can be improved by applying model interpretation techniques, such as SHAP (29) to explain the prediction made by the ML models. The other problem is that the development of the ML model requires expertise in ML and usually requires the time-consuming tuning of dozens of parameters. We have shown in the current study that this can be improved by using the recently emerging AutoML technique. AutoML can make ML model training more accessible to non-ML experts without compromise in model performance.

Comparison of an AutoML Model and Manually Derived ML Model

We have demonstrated that an AutoML derived model can achieve better performance than a manually derived ML model. The ML models need a careful selection of algorithms and tuning of hyperparameters to achieve their best performance. However, in many clinical studies that apply ML, such tuning is usually not carried out. Therefore, the developed model may not fully

exploit the power of ML. In this study, we followed the same procedures mentioned in the literature to manually develop an ML model. This represented a typical scenario in which a non-ML-expert used an open-source library to train an ML model. As a result, the manually developed model is not optimal. In contrast, AutoML can perform extensive searching of different pipelines and tuning of hyperparameters, which resulted in a better model. It has been reported that AutoML outperformed a conventional ML model manually developed by a researcher with a master's degree in computer science. Moreover, AutoML only took less than an hour to train but achieved similar or even better performance than a manually derived ML model which took days to tune (17). Another distinctive advantage of AutoML compared with the conventional ML procedure is that it is much easier to use for surgeons with limited background in ML. As we have shown, the use of the AutoML model requires only a few lines of code, which makes it more accessible to clinical doctors.

Limitations

In the current study, all the cases were from a single center and the number of cases was relatively small. Nevertheless, we have demonstrated that the use of AutoML can help clinical researchers develop high quality ML models that outperformed the statistical models and manually trained ML models. Though the current study is a single-center study with limited cases and follow-up time, the AutoML method presented in the current study can be easily generalized to a study with a larger sample size and longer follow-up time. In the current study, the treatment strategies, such as clipping, liquid embolization, or flow disruption were not assessed. To further test the applicability of our model, more cases from multiple centers with longer follow-up should be analyzed.

CONCLUSIONS

We have demonstrated the feasibility of using AutoML to develop high quality ML model for aneurysm treatment outcome prediction. The AutoML derived model accurately predicted the outcome of treatment, which may facilitate treatment planning. AutoML may outperform the conventional statistical model and manually derived machine learning model. The emergence of AutoML simplifies and automates the process of building an ML model, which lowers the learning threshold of ML and allows non-AI experts to apply ML to their research.

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/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Zhujiang Hospital of Southern Medical University. The Ethics Committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

CO completed the code, result analysis, and manuscript draft. JL completed the data preprocessing and result analysis. YQ, WC, and DL edited the manuscript. JL, XZ, and XH collected the data. XZ and C-ZD supervised the study and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation (Grants 81974177 and 81974178) and National Key Research and Development Program (Grants 2016YFC1300804 and 2016YFC1300800).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.735142/full#supplementary-material

REFERENCES

1. Liu A, Huang J. Treatment of aneurysms: clipping versus coiling. *Curr Cardiol Rep.* (2015) 17:620–28. doi: 10.1007/s11886-015-0628-2
2. Ferns SP, Sprengers ME, van Rooij WJ, Rinkel GJ, van Rijn JC, Bipat S, et al. Coiling of intracranial aneurysms: a systematic review on initial occlusion and reopening and retreatment rates. *Stroke.* (2009) 40:e523–9. doi: 10.1161/STROKEAHA.109.553099
3. Zanaty M, Chalouhi N, Starke RM, Barros G, Saigh MP, Schwartz EW, et al. Flow diversion versus conventional treatment for carotid cavernous aneurysms. *Stroke.* (2014) 45:2656–61. doi: 10.1161/STROKEAHA.114.006247
4. Fujimura S, Takao H, Suzuki T, Dahmani C, Ishibashi T, Mamori H, et al. new combined parameter predicts re-treatment for coil-embolized aneurysms: a computational fluid dynamics multivariable analysis study. *J Neurointerv Surg.* (2018) 10:791–6. doi: 10.1136/neurintsurg-2017-013433
5. Zhang Q, Jing L, Liu J, Wang K, Zhang Y, Paliwal N, et al. Predisposing factors for recanalization of cerebral aneurysms after endovascular embolization: a multivariate study. *J Neurointerv Surg.* (2018) 10:252–7. doi: 10.1136/neurintsurg-2017-013041
6. Paliwal N, Tutino VM, Shallwani H, Beecher JS, Damiano RJ, Shakir HJ, et al. Ostium ratio and neck ratio could predict the outcome of sidewall intracranial aneurysms treated with flow diverters. *Am J Neuroradiol.* (2019) 40:288–94. doi: 10.3174/ajnr.A5953
7. Piotin M, Blanc R, Spelle L, Mounayer C, Piantino R, Schmidt PJ, et al. Stent-assisted coiling of intracranial aneurysms: clinical and angiographic results in 216 consecutive aneurysms. *Stroke.* (2010) 41:110–5. doi: 10.1161/STROKEAHA.109.558114
8. Ye HW, Liu YQ, Wang QJ, Zheng T, Cui XB, Gao YY, et al. Comparison between Solitaire^{TM} AB and Enterprise stent-assisted coiling for intracranial aneurysms. *Exp Ther Med.* (2015) 10:145–53. doi: 10.3892/etm.2015.2481

9. Li H, Li XF, He XY, Zhang X, Zhu GH, Fang QR, et al. Endovascular treatment of dissecting aneurysms of the posterior inferior cerebellar artery and predictors of outcome. *J Stroke Cerebrovasc Dis.* (2015) 24:2134–42. doi: 10.1016/j.jstrokecerebrovasdis.2015.05.034
10. Paliwal N, Jaiswal P, Tutino VM, Shallwani H, Davies JM, Siddiqui AH, et al. Outcome prediction of intracranial aneurysm treatment by flow diverters using machine learning. *Neurosurg Focus.* (2018) 45:E7. doi: 10.3171/2018.8.FOCUS18332
11. O'Kelly CJ, Krings T, Fiorella D, Marotta TR. A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents. *Interv Neuroradiol.* (2010) 16:133–7. doi: 10.1177/159101991001600204
12. Kamran M, Yarnold J, Grunwald IQ, Byrne JV. Assessment of angiographic outcomes after flow diversion treatment of intracranial aneurysms: a new grading schema. *Neuroradiology.* (2011) 53:501–8. doi: 10.1007/s00234-010-0767-5
13. Raper DM, Chen CJ, Kumar J, Kalani MY, Park MS. Predicting outcomes for cerebral aneurysms treated with flow diversion: a comparison between 4 grading scales. *World Neurosurg.* (2019) 128:e209–16. doi: 10.1016/j.wneu.2019.04.099
14. Ramos LA, van der Steen WE, Barros RS, Majoie CB, van den Berg R, Verbaan D, et al. Machine learning improves prediction of delayed cerebral ischemia in patients with subarachnoid hemorrhage. *J Neurointerv Surg.* (2019) 11:497–502. doi: 10.1136/neurintsurg-2018-014258
15. Liu J, Xiong Y, Zhong M, Yang Y, Guo X, Tan X, et al. Predicting long-term outcomes after poor-grade aneurysmal subarachnoid hemorrhage using decision tree modeling. *Neurosurgery.* (2020) 87:523–9. doi: 10.1093/neuros/nyaa052
16. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med.* (2016) 375:1216. doi: 10.1056/NEJMp1606181
17. Padmanabhan M, Yuan P, Chada G, Nguyen HV. Physician-friendly machine learning: a case study with cardiovascular disease risk prediction. *J Clin Med.* (2019) 8:1050. doi: 10.3390/jcm8071050
18. Faes L, Wagner SK, Fu DJ, Liu X, Korot E, Ledsam JR, et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. *Lancet Digital Health.* (2019) 1:e232–42. doi: 10.1016/S2589-7500(19)30108-6
19. Chen Y, Fan H, He X, Guo S, Li X, He M, et al. China Intracranial Aneurysm Project (CIAP): protocol for a prospective cohort study of interventional treatment and craniotomy for unruptured aneurysms. *BMJ Open.* (2018) 8:e019333. doi: 10.1136/bmjopen-2017-019333
20. Roy D, Milot G, Raymond J. Endovascular treatment of unruptured aneurysms. *Stroke.* (2001) 32:1998–2004. doi: 10.1161/hs0901.095600
21. Trang T. Le, Weixuan Fu and Jason H. Moore Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics.* (2020) 36:250–6. doi: 10.1093/bioinformatics/btz470
22. Detmer FJ, Lückehe D, Mut F, Slawski M, Hirsch S, Bijlenga P, et al. Comparison of statistical learning approaches for cerebral aneurysm rupture assessment. *Int J Comput Assist Radiol Surg.* (2019) 15:141–50 doi: 10.1007/s11548-019-02065-2
23. Silva MA, Patel J, Kavouridis V, Gallerani T, Beers A, Chang K, et al. Machine learning models can detect aneurysm rupture and identify clinical features associated with rupture. *World Neurosurg.* (2019) 131:e46–51 doi: 10.1016/j.wneu.2019.06.231
24. Rubbert C, Patil KR, Beseoglu K, Mathys C, May R, Kaschner MG, et al. Prediction of outcome after aneurysmal subarachnoid haemorrhage using data from patient admission. *Eur Radiol.* (2018) 28:4949–58. doi: 10.1007/s00330-018-5505-0
25. Ogilvy CS, Chua MH, Fusco MR, Reddy AS, Thomas AJ. Stratification of recanalization for patients with endovascular treatment of intracranial aneurysms. *Neurosurgery.* (2015) 76:390–5. doi: 10.1227/NEU.0000000000000651
26. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE.* (2015) 10:e0118432. doi: 10.1371/journal.pone.0118432
27. Demšar J. Statistical comparisons of classifiers over multiple data sets. *J Mach Learn Res*. (2006) 7:1–30.
28. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med.* (2016) 375:1216. doi: 10.10hgd.1056/NEJMp1606181
29. 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

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.

frontiers | Frontiers in Neurology

GENERAL COMMENTARY
published: 10 June 2022
doi: 10.3389/fneur.2022.878091

Commentary: Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study

Markus Huber[1], *Markus M. Luedi*[1] *and Lukas Andereggen*[2,3]*

[1] Department of Anaesthesiology and Pain Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland, [2] Department of Neurosurgery, Kantonsspital Aarau, Aarau, Switzerland, [3] Faculty of Medicine, University of Bern, Bern, Switzerland

Keywords: AutoML, stroke, machine learning, intracranial aneurysm, endovascular treatment

OPEN ACCESS

Edited by:
Osama O. Zaidat,
Northeast Ohio Medical University,
United States

Reviewed by:
Luis Rafael Moscote-Salazar,
Latinamerican Council of Neurocritical Care (CLaNi), Colombia
Hisham Salahuddin,
Antelope Valley Hospital,
United States
Yang Wang,
Capital Medical University, China

***Correspondence:**
Lukas Andereggen
lukas.andereggen@ksa.ch
orcid.org/0000-0003-1764-688X

Specialty section:
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

Received: *17 February 2022*
Accepted: *02 May 2022*
Published: *10 June 2022*

Citation:
Huber M, Luedi MM and Andereggen L (2022) Commentary: Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study. Front. Neurol. 13:878091. doi: 10.3389/fneur.2022.878091

A Commentary on

Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study
by Ou, C., Liu, J., Qian, Y., Chong, W., Liu, D., He, X., Zhang, X., and Duan, C.-Z. (2021). Front. Neurol. 12:735142. doi: 10.3389/fneur.2021.735142

We read with great interest the article by *Ou and colleagues* (1) reporting on the application of an automated machine learning (AutoML) approach to predict recanalization after endovascular aneurysm occlusion. The authors are commended on accounting for key factors in outcome prediction such as (i) imbalanced datasets (2) by considering both the area under precision-recall curve (AUPRC) and the area under receiver-operating characteristic curve (AUROC), as well as the F1-score, (ii) the risk of overfitting by performing repeated cross-validations of the training and evaluation procedure, (iii) including graphical illustrations of the model building procedure as suggested in the literature (3) and (iv) providing code examples (4). Their results underlines the increased predictive performance of an AutoML approach compared to traditional logistic regression and a typical machine learning algorithm (Random Forest). Given the high predictive performance and the ease of using statistical software—as exemplified by the code and procedures in the Python language—the AutoML tool might provide a tool to bridge the implementation gap of such methods in medical practice (5).

From our own experience, we found the following points critical in applying ML models in outcome prediction.

While the discriminatory ability of the AutoML approach is highest among the statistical approaches in the study presented, the authors did not assess the calibration of the various algorithms. Calibration gives an estimate of how well the observed outcomes and predictions agree and are crucial in the clinical decision-making (6–8), thus we argue that an assessment of the calibration could be a further step to both evaluate and compare classical statistical methods with AutoML approaches to provide a more holistic estimate of the performance of various classifiers. As it is argued that one of the main advantages of AutoML is the possibility for non-ML experts to utilize ML models without prior know-how, we would like to point out that the application of AutoML as exemplified in the software code in Figure 4 of the paper still requires rather profound knowledge of the hyperparameters of the algorithm used in the model building pipeline—in the present application more than a dozen parameters need to be set. Thus, while the AutoML framework hides most of the parameter tuning and feature selection in a more easy-to-use

software wrapper, a certain essential knowledge of ML—such as the concept of hyperparameters and cross-validation—is still required from the user to obtain robust and unbiased results. The authors mention further drawbacks of an AutoML approach, for example in terms of the black-box problems, which could be tackled by novel interpretation techniques such as SHAP values. However, while these techniques provide information regarding the importance of individual predictors, we argue that by considering the predictive performance of an ensemble of classifiers for two performance metrics jointly provides additional valuable information to compare different algorithms (9). Thus, an illustration of the performance of various algorithms within the search for the optimal pipeline of an AutoML application might provide additional and helpful information regarding the performance and robustness of both standard statistical methods such as multivariable logistic regression and modern machine learning methods.

From a clinical perspective, recanalization and recurrences following endovascular therapy of intracranial aneurysms is not infrequently encountered. The authors indeed list the number of patients analyzed and the short-term follow-up as a study limitation. However, the short follow-up time limits its validity. Although it has been shown that coiled aneurysms that showed complete occlusion at 6 months remained stable in most cases, up to 6.5% of those aneurysm occluded completely at 6-month later showed a recanalization (10). To evaluate recurrences rates dictating the treatment effectiveness after coiling, long-term follow-up is thus warranted (11). Although a low risk of rupture of coiled aneurysms with a follow-up period of up to 20 years have been described, larger aneurysms need to be followed for a longer time period (10, 12), as do aneurysms with residual filling after the initial treatment (13). Delayed recanalization, although rare, and the possibility of de novo aneurysm formation, however calls for continuous monitoring beyond 36 months (14).

We commend the authors on presenting an interesting and important application of a novel ML approach applicable for non-AI-experts that outperforms the commonly used statistical methods in predicting treatment outcome, as the latter is of utmost importance in any clinical practice evaluating its treatment outcomes.

AUTHOR CONTRIBUTIONS

MH, MML, and LA designed and wrote the initial draft. All authors contributed to the study design and critically revised the commentary. All authors contributed to the article and approved the submitted version.

FUNDING

Open access funding was provided by the University of Bern.

REFERENCES

1. Ou C, Liu J, Qian Y, Chong W, Liu D, He X, et al. Automated machine learning model development for intracranial aneurysm treatment outcome prediction: a feasibility study. *Front Neurol.* (2021) 12:735142. doi: 10.3389/fneur.2021.735142
2. He H, Garcia EA. Learning from Imbalanced Data. *IEEE Trans Knowl Data Eng.* (2009) 21:1263–84. doi: 10.1109/TKDE.2008.239
3. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med.* (2020) 26:1320–4. doi: 10.1038/s41591-020-1041-y
4. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for reporting machine learning analyses in clinical research. *Circ Cardiovasc Qual Outcomes.* (2020) 13:e006556. doi: 10.1161/CIRCOUTCOMES.120.006556
5. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med.* (2017) 376:2507–9. doi: 10.1056/NEJMp1702071
6. Cearns M, Hahn T, Clark S, Baune BT. Machine learning probability calibration for high-risk clinical decision-making. *Aust N Z J Psychiatry.* (2020) 54:123–6. doi: 10.1177/0004867419885448
7. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* (2010) 21:128–38. doi: 10.1097/EDE.0b013e3181c30fb2
8. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* (2019) 17:230. doi: 10.1186/s12916-019-1466-7
9. Huber M, Luedi MM, Schubert GA, Musahl C, Tortora A, Frey J, et al. Machine learning for outcome prediction in first-line surgery of prolactinomas. *Front Endocrinol.* (2022) 13:810219. doi: 10.3389/fendo.2022.810219
10. Jeon JP, Cho YD, Rhim JK, Yoo DH, Kang H-S, Kim JE, et al. Extended monitoring of coiled aneurysms completely occluded at 6-month follow-up: late recanalization rate and related risk factors. *Eur Radiol.* (2016) 26:3319–26. doi: 10.1007/s00330-015-4176-3
11. Pierot L, Barbe C, Herbreteau D, Gauvrit J-Y, Januel A-C, Bala F, et al. Immediate post-operative aneurysm occlusion after endovascular treatment of intracranial aneurysms with coiling or balloon-assisted coiling in a prospective multicenter cohort of 1189 patients: Analysis of Recanalization after Endovascular Treatment of intracranial Aneurysm (ARETA) Study. *J Neurointerv Surg.* (2021) 13:918–23. doi: 10.1136/neurintsurg-2020-017012
12. Koyanagi M, Ishii A, Imamura H, Satow T, Yoshida K, Hasegawa H, et al. Long-term outcomes of coil embolization of unruptured intracranial aneurysms. *J Neurosurg JNS.* (2018) 129:1492–8. doi: 10.3171/2017.6.JNS17174
13. McDougall CG, Johnston SC, Hetts SW, Gholkar A, Barnwell SL, Vazquez Suarez JC, et al. Five-year results of randomized bioactive versus bare metal coils in the treatment of intracranial aneurysms: the Matrix and Platinum Science (MAPS) Trial. *J Neurointerv Surg.* (2021) 13:930–4. doi: 10.1136/neurintsurg-2020-016906
14. Yeon EK, Cho YD, Yoo DH, Lee SH, Kang H-S, Kim JE, et al. Is 3 years adequate for tracking completely occluded coiled aneurysms? *J Neurosur JNS.* (2020) 133:758–64. doi: 10.3171/2019.5.JNS183651

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.

frontiers in Neurology

MINI REVIEW
published: 23 February 2022
doi: 10.3389/fneur.2022.784326

Role of Artificial Intelligence in Unruptured Intracranial Aneurysm: An Overview

Anurag Marasini[1], *Alisha Shrestha*[2], *Subash Phuyal*[3], *Osama O. Zaidat*[4] and *Junaid Siddiq Kalia*[1,5,6]*

[1] AINeuroCare Academy, Dallas, TX, United States, [2] Travel and Mountain Medicine Center, Kathmandu, Nepal, [3] Department of Neurointerventional Radiology, Upendra Devkota Memorial National Institute of Neurological and Allied Sciences, Kathmandu, Nepal, [4] Departments of Endovascular Neurosurgery and Stroke, St. Vincent Mercy Medical Center, Toledo, OH, United States, [5] Clinical Strategy, VeeOne Health Inc., Roseville, CA, United States, [6] neurologypocketbook.com, Dallas, TX, United States

Intracranial aneurysms (IAs) are a significant public health concern. In populations without comorbidity and a mean age of 50 years, their prevalence is up to 3.2%. An efficient method for identifying subjects at high risk of an IA is warranted to provide adequate radiological screening guidelines and effectively allocate medical resources. Artificial intelligence (AI) has received worldwide attention for its impressive performance in image-based tasks. It can serve as an adjunct to physicians in clinical settings, improving diagnostic accuracy while reducing physicians' workload. AI can perform tasks such as pattern recognition, object identification, and problem resolution with human-like intelligence. Based on the data collected for training, AI can assist in decisions in a semi-autonomous manner. Similarly, AI can identify a likely diagnosis and also, select a suitable treatment based on health records or imaging data without any explicit programming (instruction set). Aneurysm rupture prediction is the holy grail of prediction modeling. AI can significantly improve rupture prediction, saving lives and limbs in the process. Nowadays, deep learning (DL) has shown significant potential in accurately detecting lesions on medical imaging and has reached, or perhaps surpassed, an expert-level of diagnosis. This is the first step to accurately diagnose UIAs with increased computational radiomicis. This will not only allow diagnosis but also suggest a treatment course. In the future, we will see an increasing role of AI in both the diagnosis and management of IAs.

Keywords: artificial intelligence, machine learning, aneurysm rupture, unruptured intracranial aneurysms (UIAs), deep learning

OPEN ACCESS

Edited by:
Shakir Husain Hakim, University Hospital Zürich, Switzerland

Reviewed by:
Sandrine de Ribaupierre, Western University, Canada
S. Ottavio Tomasi, Paracelsus Medical University, Austria

***Correspondence:**
Junaid Siddiq Kalia
junaidkalia@aineurocare.com

Specialty section:
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

Received: *27 September 2021*
Accepted: *04 January 2022*
Published: *23 February 2022*

Citation:
Marasini A, Shrestha A, Phuyal S, Zaidat OO and Kalia JS (2022) Role of Artificial Intelligence in Unruptured Intracranial Aneurysm: An Overview. Front. Neurol. 13:784326. doi: 10.3389/fneur.2022.784326

INTRODUCTION

Intracranial aneurysms (IAs) are a significant public health concern. In populations without comorbidity and a mean age of 50 years, their prevalence is up to 3.2% (1). With the development and application of advanced imaging techniques worldwide, unruptured intracranial aneurysms (UIAs) are being detected more frequently. Subarachnoid hemorrhage (SAH) owing to UIA rupture accounts for 5–10% of all strokes in the United States (2) SAH may cause considerably high mortality, and those who survive may endure chronic neuropsychological effects and decreased quality of life (3). Whether detected incidentally or during screening, the management of intracranial aneurysms (IA) is a challenge in itself for both the treating physician and patients.

This challenge cannot be overcome easily by conventional methodology because the decision to intervene involves weighing the periprocedural risks innate to endovascular or surgical methods and subarachnoid hemorrhage (4). Due to a lack of clear understanding of the natural history of small UIAs and the heterogeneity in the current guidelines and literature, there is significant variability in the perceptions and surveillance practices for small UIAs (<7 mm). This was discussed in a survey of 227 practicing neuroradiologists and neurointerventionalists. Where there is clear variability in practice both in terms of frequency and method of follow-up imaging. About 59% favored indefinite, life-long follow-up for small unruptured intracranial aneurysms, and favored non-contrast MR angiography for aneurysm follow-up (5).

This heterogeneity and variability in practice necessitate tools and models that can improve overall recommendations. More recently AI has started to gain attention from researchers and clinicians alike for this problem. Numerous researchers have been exploring artificial intelligence (AI) and Deep Learning's (DL) ability in the field of aneurysm detection, rupture prediction, and also improvement of workflow. AI can perform tasks such as pattern recognition, object identification, and problem resolution with human-like intelligence. Based on the data collected for training, AI can make autonomous decisions. Similarly, in medicine, AI can recognize a likely diagnosis and select a suitable treatment based on health records or imaging data without any explicit programming. Machine learning (ML) endows AI to learn and train models to extract and memorize features and related parameters. There are three types of ML: supervised (training with specific labels or annotations), unsupervised (training without specific labels and the algorithm clusters data to reveal underlying patterns), and semi- or weakly supervised learning (training with both labeled and unlabeled data to reduce the annotation burden) (4).

Nowadays, deep learning (DL) has shown significant potential in accurately detecting lesions on medical imaging and has reached, or perhaps surpassed, an expert-level of diagnosis (6). DL is a machine learning technique that directly learns the most predictive features from a large data set of labeled images (6). A deeper discussion on deep learning is beyond the scope of this mini-review. Many introductions are written on artificial intelligence; we recommend reading "High-performance medicine: the convergence of human and artificial intelligence" by Topol et al. (7). We also recommend Liu et al. (8) guide to read articles that use machine learning. Briefly, the technique primarily used in neuroimaging is convolutional neural networks (CNN), a subset of Deep Learning. The visual cortex inspires the algorithms based on CNN's which closely resemble human neurons. They are used extensively in image feature selection, classification, and segmentation, etc., (9).

AI IN INTRACRANIAL ANEURYSM DETECTION

CT-Angiography and MR-Angiography have been used widely to screen for IAs. Manual detection of pathologies from images is a laborious process and requires 3D modeling to accurately depict vessel morphology. The process is time-consuming and slow. With the advent of computer vision and deep learning, we can have AI directly analyze images for the presence of IAs. However, with newer imaging techniques and newer models, we gradually saw the AI performance increase compared to human counterparts and at times surpass it. This directly improves the detection rate. As earlier detection allows closer monitoring and eventually results in better patient selection for intervention.

In 2020, Zhao et al. (10) compared the performance of their 3D CNN segmentation model DAResUNET trained on 1,117 Digital Subtraction Angiography (DSA) images against six board-certified radiologists and two neurosurgeons. For aneurysms > 5 mm in size, the ML model had a superior accuracy of 100%. However, four aneurysms < 3 mm in size were missed resulting in a sensitivity of 98.6% for aneurysms < 3 mm in size (10).

Aneurysms can be missed due to the small size and complexity of intracranial vessels. AI can be used to overcome this complexity; as shown by Yuki et al. In their case series, their AI model ResNet-18 detected 5 aneurysms < 2 mm in size by using maximum intensity projection (MIP), to analyze the image in different rotational views. Unfortunately, two board-certified radiologists initially had missed the same five aneurysms. While a human operator can analyze different images, it is time-consuming. ML models can assist in improving accuracy as now we have higher computational power, thus decreasing the latency in detection as these models can predict in near real-time (11).

This was rightly demonstrated by Yang et al. CNN-based algorithm that detected cerebral aneurysms with a sensitivity of 97.5% among 649 computational tomography (CT) images. It also detected eight new aneurysms that were missed in the initial reports, thereby improving the overall performance of radiologists in terms of area under weighted alternative free-response receiver operating characteristic from 0.60 to 0.61. The study showed that AI has good potential when it comes to being a supportive tool for radiologists rather than a replacement (12).

This potential was further explored by Faron et al. (13) who compared the performance of a DEEPMedic CNN model to detect aneurysms from clinical TOF-MRA data vs. two expert human readers. No statistically significant difference was found in the overall sensitivity (OS), showing that AI has now started to reach human-level accuracy. However, when the human reader and CNN's detections were combined, the OS of both human readers was improved (reader 1: 98 vs. 95%, $P = 0.280$; reader 2: 97 vs. 94%, $P = 0.333$). In fact, four previously undetected aneurysms were detected with this combination, reinforcing the fact that AI will continue to improve our diagnostic capabilities and definitely improve patient outcomes (13).

AI IN INTRACRANIAL ANEURYSM SCREENING

The burden of disability and mortality from UIAs is significant, while routine screening is challenging because of financial resources, logistical resources, contrast, and radiation load.

Moreover, the growth of aneurysms can be non-linear, so the timing of follow-up scans needs to be further personalized. Health insurance-related implications can also arise due to the detection of an IA. We need better guidelines for targeted screening of high-risk individuals and more sophisticated tools to time follow-up studies. These need to be evidence-based tools. AI-based increase in detection of UIAs will help us improve detection on individual patients and improve our detection strategy at a population level.

AI and ML can help us better identify screening targets. Current UIA screening guidelines in the United States and Korea are limited to two categories: (1) patients with at least two family members with UIA or SAH, and (2) patients with a history of autosomal dominant polycystic kidney disease (ADPKD), coarctation of the aorta, or microcephalic osteodysplastic primordial dwarfism. Heo et al. (14) extracted data from the National Health Screening Program in Korea containing general health examinations from 2009 to 2013. Using 21 variables from this data, Logistic Regression (LR), Random Forest (RF), eXtreme Gradient Boosting (XGB), and Deep Neural Network (DNN) were trained, among which the highest area under receiver operating curve (AUROC) value was achieved by the XGB algorithm (0.765) (95% CI 0.742–0.788). The authors stratified the risk group into five categories. This risk stratification with the help of AI models will help in improved targets for screening. This targeted screening in the future with the use of multimodal data: health status & history, family and similar population imaging analysis, genetics will eventually create new guidelines for us to follow (14).

AI IN INTRACRANIAL ANEURYSM RUPTURE PREDICTION

Rupture prediction is the holy grail of prediction modeling. AI can significantly improve rupture prediction, saving lives and limbs in the process. Conventional logistic regression (LR) has been one of the most studied statistical models to predict the rupture status of unruptured intracranial aneurysms (UIAs). In recent years, numerous studies have been done to develop and compare the utility of ML in the prediction of aneurysm rupture risk in the context of statistical models.

Multimodal data is needed for ML to improve its accuracy. Data only limited to imaging does not perform as well in prediction modeling. Chen et al. concluded that integrating clinical, aneurysm morphological, and hemodynamic parameters improves rupture prediction (15).

Similarly, Detmer et al. also compared the logistic regression probability model (LRM) to other ML classifiers by training both the models with hemodynamic, morphological, and patient-related information of 1,631 intracranial aneurysms. The predictive performance of ML classifiers was comparable to the group lasso model. They concluded that incorporating additional information such as aneurysm vessel wall enhancement would lead to better performance of the ML classifiers (16).

We have used scores in clinical practice to predict rupture—one of the widely used includes the PHASES Score (17). Zhu et al. compared ML models with statistical methods and PHASES score in intracranial aneurysm stability assessment. Among the authors' three ML models [Artificial Neural Network (ANN), RF and Support Vector Machine (SVM)], ANN showed the best performance with an area under the curve (AUC) of 0.851 (95% CI 0.828–0.873). Interestingly, even the least performing ML model, RF (0.832 (95% CI 0.809–0.855) significantly outperformed the statistical models and the PHASES score ($P = 0.045$ and $P < 0.001$, respectively). Thus implying that ML models provide better accuracy when compared to commonly used statistical tools such as logistic regression (LR). The superiority of ML over traditional statistical methods can be attributed to the fact that ML has the capacity to simultaneously process massive numbers of variables and can model non-linear relationships while LR and PHASES are limited to linear relationships (18).

AI IN INTRACRANIAL ANEURYSM IN CLINICAL DECISION SUPPORT

Therapeutic planning after detection of UIA is very complex and depends on patient factors as well as the characteristics of the aneurysm. Regarding the aneurysm, its location, size, and feeding artery are a few parameters. Meanwhile, a neurovascular multidisciplinary team must be involved to analyze the risk of aneurysmal rupture, the risks of endovascular and surgical treatment, and the predicted outcome of treatment taking into account factors such as the patient's age, lifestyle, comorbid conditions, and personal preferences (19).

Thus AI models that can analyze multiple parameters simultaneously and work with large volumes of data can aid this complex decision making. In addition, AI can enhance the process by adding additional objective data of flow and morphological characteristics of aneurysms. This will lead to an improved occlusion rate of aneurysms and potentially decreased recanalization rates.

In 2019, Bhurwani et al. (20) developed a DNN using Keras to predict occlusion treatment outcomes as a binary output: occluded or unoccluded using only intraoperative information. They analyzed 190 CAs pre and post. This feasibility study concluded with quantitative imaging information that is normalized and improves prediction. Also, shown parameters at an individual level can improve accuracy (20).

As previously mentioned, we have used statistical modeling to develop clinical scores to predict occlusion rates. One such score DIANES score is being used. To predict occlusion success, they used six features (IA diameter, indication, parent artery diameter ratio, neck ratio, side-branch artery, and sex) (21). Williams et al. (22) developed Aneurysm Occlusion Assistant (AnOA), a platform approach is envisioned to assist in real-time decision making. They developed models that can more reliably help with therapeutic decision-making. They have envisioned a new platform that would be able to bring the analytical frameworks from the lab into clinical settings to guide real-time decision

making; the Aneurysm Occlusion Assistant (AnOA), an AI assistant based on Keras, Tensorflow, and skLearn that aims to assist neurosurgeons intra-operationally in order to personalize endovascular aneurysm treatments better. Although this has not been tested clinically, the system used pre-and post-device placement data as input and allowed for segmentation of IAs and cranial vasculature with a dice index of ∼0.78 and was able to predict aneurysm occlusion at 6 months with accuracy 0.84, in 6.88 (22).

This exemplifies the future of AI in clinical decision-making, and these platforms will start integrating as clinical evidence is accumulated. We will see similar approaches being applied to therapeutic planning as we have already seen in the field of cardiology (23, 24). In the future, we will see these models tested in real-world scenarios and give real-time automatic suggestions for therapeutic planning.

DOCUMENTATION, QUALITY, AND WORKFLOW ENHANCEMENT USING AI

There is a significant shortage of neuroradiologists, especially in the developing world. In addition, there is an increasing burden of newer forms of imaging data, including 2D and 3D imaging, even in the developed world. AI can augment radiologists in their workflow by decreasing latency, improving accuracy, and automatically performing repetitive tasks like measurements to save time. Dai et al. used modified 2D CTA images produced by the nearby projection (NP) method to train a fast Region-Based Convolutional Neural Network (RCNN) model. The model automatically proposed rectangular areas on the images that may contain aneurysms so that radiologists can easily find them and check whether these proposals were correct or not. The model was 91.8% sensitive in detecting aneurysms automatically. While an experienced radiologist takes around 10 to 15 min to complete one aneurysm image diagnosis, the authors discuss that with their model, it will only take a few minutes for a radiologist to observe all the proposals so that they can save almost 10 min per case, thereby improving efficiency and significantly reducing detection time.

One of the more explored aspects of AI in improving workflow has to be automated note creation using speech recognition software. This has allowed physicians to save time as machine learning converts speech into text (25). An intelligent Electronic Medical Record (EMR) can generate an automated summary resulting in a timely wrap-up of the care process. In addition, AI algorithms can be integrated into various steps in patient management, thus streamlining this critical process. Recently, Williams et al. have reported one of the first papers discussing Aneurysm Occlusion Assistant (AnOA), a semi-autonomous system that uses AI, which can predict the surgical outcome of an IA immediately following device placement, allowing for therapy adjustment. This study marks a crucial step toward incorporating machine learning and data-driven algorithms into surgical suites for better treatment outcomes.

CHALLENGES AND LIMITATIONS OF AI IN INTRACRANIAL ANEURYSM

Despite significant advancements, this technology is still in its infancy, and before deploying it in a clinical setting, it needs to be thoroughly tested. Given that the majority of data will be imaging-based, we will need faster bandwidth and more processing with dedicated hardware built from the ground up dedicated to processing AI-based tasks. Newer computer chips are now aggressively integrating AI at the level of chip design (26).

The continued biggest challenge in non-imaging data (namely EHR-based data) remains a considerable bottleneck in innovation in healthcare in general and AI in particular. We need more and better standardization of data, improved data sharing, and API-based integrations. The new Cures act has addressed a few of these issues, but we still have a long way to go (27, 28).

CONCLUSION

Recent evidence shows that AI, especially Deep Learning, is evolving as a promising aide in clinical decision making in medicine. AI grants us the computational power to explore complicated non-linear relationships in extensive amounts of data, and its predictive power increases with the available datasets for training. Thus the massive amount of data accumulating in clinics, hospitals and stored in electronic medical records through standard tests and medical imaging allows for more applications of AI and high-performance data-driven medicine. With the need for well-trained radiologists and the amount of imaging data generated in healthcare settings worldwide, AI-based CADs will be a tool that will help neuroradiologists streamline clinical workflow while approaching clinical problem solving efficiently and accurately. Our review explored the frontier on how AI can detect aneurysms, evaluate rupture risk, help in triaging clinical therapy strategies, predict treatment outcomes and enhance workflow. Although we have not quite yet reached the threshold for routine clinical application, we believe that with the availability of larger datasets, AI has great potential to solve intracranial aneurysm management issues in a patient-centric manner. Evidence suggests that AI models have started to match and even outperformed human readers on numerous occasions while interpreting medical images. Thus, it would not be an understatement to say that an AI-powered real-time decision-making assistant software for clinics, hospitals, and operating suites will be a norm in the coming years. Artificial Intelligence in neuroradiology; the future is already here.

A GLIMPSE INTO THE FUTURE-AI-ENHANCED INTRACRANIAL ANEURYSM CARE

A 24-year-old female had a right-sided pounding headache with left-sided weakness, numbness, and tingling. She was rushed to the ER with concern for acute ischemic stroke however her final diagnosis was hemiplegic migraine. A CTA was done in

accordance with the standard of care for acute stroke. The neurologist focused on stroke and then migraine. However, a 4 mm aneurysm was also detected in her internal carotid artery with the help of an AI-based system and recommended intervention. The system had evaluated patient rupture risk not just on the size as previously thought to be the main criteria. Rather using multimodal data including imaging; feeding artery, diameters, ratios, radiomic; flow mechanics, genomics, and metabolomics. On the day of the procedure, an AI-based software analyzed real-time images of digital subtraction angiography (DSA) and recommended the type and size of the stent to be used. Post-procedure the appropriate choice of antithrombotics was assisted by AI according to information of her genomics on file. This added case went to a central registry which helped identify two of her relatives for screening. One of whom underwent an aneurysm occlusion procedure for her undiagnosed 8 mm aneurysm.

AUTHOR CONTRIBUTIONS

AM: did initial systemic search to find articles, then proceeded to with progressive summarization, and authored majority of the article with JK. AS: did initial systemic search to find articles and then proceeded with progressive summarization. SP: independent search and helped with current and future clinical relevance of said articles. OZ: reviewed and summarized articles for accuracy and relevance. JK: contributed in all aspects of the manuscript including editing and writing the final manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol.* (2011) 10:626–36. doi: 10.1016/S1474-4422(11)70109-0
2. Rincon F, Rossenwasser RH, Dumont A. The epidemiology of admissions of nontraumatic subarachnoid hemorrhage in the United States. *Neurosurgery.* (2013) 73:217–23. doi: 10.1227/01.neu.0000430290.93304.33
3. Ois A, Vivas E, Figueras-Aguirre G, Guimaraens L, Cuadrado-Godia E, Avellaneda C, et al. Misdiagnosis worsens prognosis in subarachnoid hemorrhage with good hunt and hess score. *Stroke.* (2019) 50:3072–6. doi: 10.1161/STROKEAHA.119.025520
4. Zhou Z-H. A brief introduction to weakly supervised learning. *Natl Sci Rev.* (2017) 5:44–53. doi: 10.1093/nsr/nwx106
5. Malhotra A, Wu X, Geng B, Hersey D, Gandhi D, Sanelli P. Management of Small Unruptured Intracranial Aneurysms: A Survey of Neuroradiologists. *American Journal of Neuroradiology.* (2018) 39:875–80. doi: 10.3174/ajnr.A5631
6. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* (2019) 25:44–56. doi: 10.1038/s41591-018-0300-7
8. Liu Y, Chen P-HC, Krause J, Peng L. How to read articles that use machine learning. *JAMA.* (2019) 322:1806. doi: 10.1001/jama.2019.16489
9. Suzuki K. Overview of deep learning in medical imaging. *Radiol Phys Technol.* (2017) 10:257–73. doi: 10.1007/s12194-017-0406-5
10. Shi Z, Miao C, Schoepf UJ, Savage RH, Dargis DM, Pan C, et al. A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. *Nat Commun.* (2020) 11:6090. doi: 10.1038/s41467-020-19527-w
11. Shimada Y, Tanimoto T, Nishimori M, Choppin A, Meir A, Ozaki A, et al. Incidental cerebral aneurysms detected by a computer-assisted detection system based on artificial intelligence. *Medicine.* (2020) 99:e21518. doi: 10.1097/MD.0000000000021518
12. Yang J, Xie M, Hu C, Alwalid O, Xu Y, Liu J, et al. Deep learning for detecting cerebral aneurysms with CT angiography. *Radiology.* (2021) 298:155–63. doi: 10.1148/radiol.2020192154
13. Faron A, Sichtermann T, Teichert N, Luetkens JA, Keulers A, Nikoubashman O, et al. Performance of a deep-learning neural network to detect intracranial aneurysms from 3D TOF-MRA compared to human readers. *Clin Neuroradiol.* (2019) 30:591–8. doi: 10.1007/s00062-019-00809-w
14. Heo J, Park SJ, Kang S-H, Oh CW, Bang JS, Kim T. Prediction of intracranial aneurysm risk using machine learning. *Sci Rep.* (2020) 10:6921. doi: 10.1038/s41598-020-63906-8
15. Chen G, Lu M, Shi Z, Xia S, Ren Y, Liu Z, et al. Development and validation of machine learning prediction model based on computed tomography angiography–derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. *Eur Radiol.* (2020) 30:5170–82. doi: 10.1007/s00330-020-06886-7
16. Detmer FJ, Lückehe D, Mut F, Slawski M, Hirsch S, Bijlenga P, et al. Comparison of statistical learning approaches for cerebral aneurysm rupture assessment. *Int J Comput Assist Radiol Surg.* (2019) 15:141–50. doi: 10.1007/s11548-019-02065-2
17. Greving JP, Wermer MJH, Brown RD, Morita A, Juvela S, Yonekura M, et al. Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies. *Lancet Neurol.* (2014) 13:59–66. doi: 10.1016/S1474-4422(13)70263-1
18. Zhu W, Li W, Tian Z, Zhang Y, Wang K, Zhang Y, et al. Stability assessment of intracranial aneurysms using machine learning based on clinical and morphological features. *Transl Stroke Res.* (2020) 11:1287–95. doi: 10.1007/s12975-020-00811-2
19. Renowden S, Nelson R. Management of incidental unruptured intracranial aneurysms. *Pract Neurol.* (2020) 20:347–55. doi: 10.1136/practneurol-2020-002521
20. Shiraz Bhurwani MM, Waqas M, Podgorsak AR, Williams KA, Davies JM, Snyder K, et al. Feasibility study for use of angiographic parametric imaging and deep neural networks for intracranial aneurysm occlusion prediction. *J NeuroIntervent Surg.* (2019) 12:714–9. doi: 10.1136/neurintsurg-2019-015544
21. Guédon A, Thépenier C, Shotar E, Gabrieli J, Mathon B, Premat K, et al. Predictive score for complete occlusion of intracranial aneurysms treated by flow-diverter stents using machine learning. *J Neurointerv Surg.* (2020) 13:341–6. doi: 10.1136/neurintsurg-2020-016748
22. Williams KA, Podgorsak AR, Shiraz Bhurwani MM, Rava RA, Sommer KN, Ionita CN. The aneurysm occlusion assistant, an AI platform for real time surgical guidance of intracranial aneurysms. *Proc SPIE Int Soc Opt Eng.* (2021) 11601:116010V. doi: 10.1117/12.2581003
23. *London's Royal Free Hospital to Use AI keyhole Procedure For Heart Attack Patients.* MobiHealthNews (2021). Available online at: https://www.mobihealthnews.com/news/emea/londons-royal-free-hospital-use-ai-keyhole-procedure-heart-attack-patients (accessed September 27, 2021).
24. *UltreonTM 1.0 Software for OCT Intravascular Imaging Abbott.* Cardiovascularabbott (2021). Available online at: https://www.cardiovascular.abbott/int/en/hcp/products/percutaneous-coronary-intervention/intravascular-imaging/ultreon-software/about.html (accessed September 27, 2021).
25. Sindhu Kutty. *Council Post: The Rise Of AI Voice Assistants In Clinical Documentation. Forbes* (2021). Available online at: https://www.forbes.com/sites/forbesbusinesscouncil/2021/03/03/the-rise-of-ai-voice-assistants-in-clinical-documentation/?sh=6bcf449356cc (accessed September 27, 2021).
26. Kahng AB. AI system outperforms humans in designing floorplans for microchips. *Nature.* (2021) 594:183–5. doi: 10.1038/d41586-021-01515-9

27. Office of the Commissioner. *21st Century Cures Act. US Food and Drug Administration*. (2020). Available online at: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act (accessed September 27, 2021).
28. ONC's Cures Act Final Rule. *Healthitgov*. (2021). Available online at: https://www.healthit.gov/curesrule/overview/about-oncs-cures-act-final-rule (accessed September 27, 2021).

Conflict of Interest: JK is employed by VeeOne Health.

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.

frontiers | Frontiers in Neurology

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fneur.2022.884693

Interpretable Machine Learning Modeling for Ischemic Stroke Outcome Prediction

Mohamed Sobhi Jabal[1], Olivier Joly[2], David Kallmes[1], George Harston[2,3], Alejandro Rabinstein[4], Thien Huynh[5] and Waleed Brinjikji[1]*

[1] Department of Radiology, Mayo Clinic, Rochester, MN, United States, [2] Brainomix Limited, Oxford, United Kingdom, [3] Oxford University Hospitals National Health Service Trust, Oxford, United Kingdom, [4] Department of Neurology, Mayo Clinic, Rochester, MN, United States, [5] Department of Radiology, Mayo Clinic, Jacksonville, FL, United States

Background and Purpose: Mechanical thrombectomy greatly improves stroke outcomes. Nonetheless, some patients fall short of full recovery despite good reperfusion. The purpose of this study was to develop machine learning (ML) models for the pre-interventional prediction of functional outcome at 3 months of thrombectomy in acute ischemic stroke (AIS), using clinical and auto-extractable radiological information consistently available upon first emergency evaluation.

Materials and Methods: A two-center retrospective cohort of 293 patients with AIS who underwent thrombectomy was analyzed. ML models were developed to predict dichotomized modified Rankin score at 90 days (mRS-90) using clinical and imaging features, both separately and combined. Conventional and experimental imaging biomarkers were quantified using automated image-processing software from non-contract computed tomography (CT) and computed tomography angiography (CTA). Shapley Additive Explanation (SHAP) was applied for model interpretability and predictor importance analysis of the optimal model.

Results: Merging clinical and imaging features returned the best results for mRS-90 prediction. The best performing classifier was Extreme Gradient Boosting (XGB) with an area under the receiver operating characteristic curve (AUC) = 84% using selected features. The most important classifying features were age, baseline National Institutes of Health Stroke Scale (NIHSS), occlusion side, degree of brain atrophy [primarily represented by cortical cerebrospinal fluid (CSF) volume and lateral ventricle volume], early ischemic core [primarily represented by e-Alberta Stroke Program Early CT Score (ASPECTS)], and collateral circulation deficit volume on CTA.

Conclusion: Machine learning that is applied to quantifiable image features from CT and CTA alongside basic clinical characteristics constitutes a promising automated method in the pre-interventional prediction of stroke prognosis. Interpretable models allow for exploring which initial features contribute the most to post-thrombectomy outcome prediction overall and for each individual patient outcome.

Keywords: ischemic stroke, artificial intelligence, machine learning, prognosis, prediction model

OPEN ACCESS

Edited by:
Osama O. Zaidat, Northeast Ohio Medical University, United States

Reviewed by:
Franziska Dorn, University Hospital Bonn, Germany
Niko Sillanpää, Tampere University Hospital, Finland

****Correspondence:***
Mohamed Sobhi Jabal
jabal.mohamedsobhi@mayo.edu

Specialty section:
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

Received: *26 February 2022*
Accepted: *20 April 2022*
Published: *19 May 2022*

Citation:
Jabal MS, Joly O, Kallmes D, Harston G, Rabinstein A, Huynh T and Brinjikji W (2022) Interpretable Machine Learning Modeling for Ischemic Stroke Outcome Prediction. Front. Neurol. 13:884693. doi: 10.3389/fneur.2022.884693

INTRODUCTION

Mechanical thrombectomy is currently the standard of care for patients with disabling the stroke from large vessel occlusion. Numerous trials have demonstrated its efficacy in improving survival and functional outcome for these patients (1–3). In addition, successful reperfusion does not translate into favorable recovery for a substantial proportion of patients who are treated with mechanical thrombectomy (4, 5). Thus, accurate and time-efficient risk assessment remains crucial to optimize triaging and outcomes of patients who may be candidates for stroke reperfusion therapy (6).

Several prognostication scales have been proposed to predict the functional outcome of patients with ischemic stroke (7–9). However, these scores lack the ability to fully model the complex and non-linear relationships between various prognostic factors with functional outcomes and they depend mainly on categorical rendering of clinical and few conventional imaging features.

Machine learning (ML) has emerged as a promising tool for fitting and modeling complex and multidimensional data patterns, leading to many potential applications in Medicine (10). This is owing to its ability to incorporate a large number of variables, extract nuanced information, and generalize the acquired knowledge on new unseen cases in an efficient and automatic manner, which could be particularly helpful in time-critical situations, such as acute stroke. Artificial intelligence (AI) algorithms could help to improve prediction methods by providing immediate prognostic information. Few studies have applied ML models on multimodal imaging features for modified Rankin Score (mRS) prediction in ischemic stroke, however, there still is ample room for refinement (11–13).

An important challenge for AI applications in healthcare is to overcome the confidence barrier and ensure that physicians trust the ensuing results. The black-box nature of ML algorithms makes it difficult to interpret most complex models, significant progress though has been made in the last few years in ML interpretability. One particularly promising method is Shapley Additive Explanations (SHAP), which to our knowledge has not been previously explored in depth for ML prediction of functional recovery after ischemic stroke.

The aim of this study was to develop an ML model and assess its potential in pre-interventional prediction of functional outcomes at 3 months of thrombectomy in acute ischemic stroke (AIS) using clinical and auto-extractable radiological information consistently available upon first evaluation in the emergency department. In addition, to establish an automated end-to-end system for streamlined patient triage and management decision support in stroke.

Abbreviations: AIS, Acute Ischemic Stroke; ASPECTS, Alberta Stroke Program Early CT Score; AUC, Area Under the Receiver Operating Characteristic Curve; AI, Artificial Intelligence; CSF, Cerebrospinal Fluid; CT, Computed Tomography; CTA, Computed Tomography Angiography; CTA-CS, CTA Collateral Score; XGB, Extreme Gradient Boosting; GB, Gradient Boosting; IQR, Interquartile Range; KNN, K Nearest Neighbor; ML, Machine Learning; mRS, modified Rankin Score; mRS-90, modified Rankin Score At 90 Days; NIHSS, National Institutes of Health Stroke Scale; RF, Random Forests; SHAP, Shapley Additive Explanations; TICI, Treatment in cerebral infarction.

MATERIALS AND METHODS

Dataset

The study included 443 patients from two academic centers with a confirmed diagnosis of AIS, due to large vessel occlusion in the anterior circulation [internal carotid or middle cerebral artery (MCA)] confirmed on computed tomography angiography (CTA) who underwent mechanical thrombectomy between 2014 and 2020. All included patients underwent standardized acute stroke imaging that includes non-contrast head CT and CTA of the head and neck. Patients without CTA or with any missing pertinent clinical or radiological data were excluded. The primary clinical outcome of interest was the modified Rankin score at 90 days (mRS-90). The study protocol was approved by the Institutional Review Board. The data that support the findings of this study could be available from the corresponding author upon reasonable request.

Feature Extraction

Collected clinical and demographic characteristics were age, sex, baseline National Institutes of Health (NIH) Stroke Scale (NIHSS), time from symptom onset to admission, and comorbidities (diabetes, hypertension, hyperlipidemia, previous stroke, and cardiovascular disease, such as myocardial infarction and arrhythmia), in addition to blood glucose and blood pressure levels. Interventional and post-interventional features, such as modified treatment in cerebral infarction (mTICI) score informing reperfusion status, were excluded from the models in line with the main purpose of the study.

Quantitative imaging feature extraction was performed using e-Stroke software (Brainomix, Oxford, UK) for automated calculation of Alberta Stroke Program Early CT Score (ASPECTS; e-ASPECTS) and estimated acute infarct volumes on non-contrast CT (14–17). e-ASPECTS uses ML classification to distinguish and segment regions that contain signs consistent with the acute ischemic change in order to output both (total and per ASPECTS region) and total e-ASPECTS volumes. Additional novel features that were extracted using e-Stroke software included non-acute infarct volume, total brain volume, and atrophy, which were quantified using cerebrospinal fluid (CSF) segmentation volumes in both the lateral ventricles and the cortical sulci separately and expressed as percentages. e-CTA (Brainomix, Oxford, UK) identifies large vessel occlusion site and quantifies the volume of collateral circulation deficit both as a percentage of the total volume and using the CTA collateral score (CTA-CS) (18–20). Novel experimental outputs from e-CTA included the absolute volume of the vessel density deficit in MCA territory relative to the contralateral hemisphere.

Feature Pre-processing

Baseline features were categorized into clinical and imaging feature groups. Standardization scaling of continuous and ordinal feature values was applied to obtain a mean of zero and a standard deviation (SD) of 1, in order to facilitate the algorithm learning process and improve the prediction results. Random splitting of the datasets into a training set and a testing set was applied with a ratio of 75–25%, respectively. The

mRS-90 was dichotomized with mRS 0–2 representing a good functional outcome.

The features were divided into four subsets, which are as follows: (1) clinical features, (2) imaging features, (3) combined clinical and imaging features, and (4) selected features. A model-based approach was applied using sequential backward feature selection with a bagging classifier, where an algorithm sequentially removes features from the full feature set until the removal of further features decreases the classifier performance.

Statistical Analysis

Statistical assessment of each clinical and image-based feature in relation to mRS-90 was assessed using the chi-square test for the categorical variables, Wilcoxon Rank-Sum test, and *t*-test for the ordinal and continuous variables depending on the normality of their distributions. Statistical analysis was done using Python (version 3.9) and the SciPy library. Values of $p < 0.05$ were considered statistically significant.

ML Model Development and Testing

For the purpose of mRS-90 prediction, supervised ML classification methods were deployed. The ML algorithms used were as follows: k-nearest neighbors, random forests (RF), gradient boosting (GB), and Extreme Gradient Boosting (XGBoost). The models were constructed using the Scikit-learn library.

As a first step, *k*-fold cross-validation of 10-folds was performed during the training for each model, which divides the training set into 10 subsets (9 for training and 1 for validation), where the training and validation sets change and iterate over the 10-folds. The model hyperparameters were optimized by means of a grid search approach, where for every model and for each hyperparameter a set of possible values was manually defined and evaluated exhaustively in every iteration to determine the values corresponding to the model's highest performance, with an area under the receiver operating characteristic curve (AUC) as scoring metric. The ML models were trained using each of the 4 different feature categories. Subsequently, for every feature group, we tested the models' performance on the testing set of patients.

From the output of the grid search, the best performing model was chosen. Finally, automatic Bayesian hyperparameter tuning with the Optuna framework was used on the best performing model to boost its performance and achieve finer tuning. The evaluation metrics used were accuracy, F1 score (for mRS-90 ≤ 2 and mRS-90 > 2 predictions), and AUC.

To enhance the model's explainability and perform a feature importance analysis, we used the method SHAP, which is based on game theory and consists of computing Shapley values reflecting the contribution of each feature in the predictions of the model (21–23). The method allows for the identification of features with the most influence on model output and measures the impact if each variable was to be removed while taking into account the interaction with other variables that provide insight on the relative importance of the features used by the model for its prediction decision process.

An illustrative summary of the methods is provided in **Figure 1**.

RESULTS

Patient Population

Of 443 total patients (266 from the first center and 177 from the second center), 293 patients met the study inclusion criteria. The remainder of the patients were excluded for lacking relevant variables mainly CTA images. In total, 101 patients had a favorable functional outcome (mRS-90 ≤ 2), while 192 patients had unfavorable functional outcomes (mRS-90 > 2). The median age of included patients was 71 years and 49% ($n = 143$) were women (**Table 1**).

Univariate Statistical Analysis

Favorable clinical outcome was significantly associated with younger age ($p < 0.0001$), female sex ($p = 0.016$), and lower baseline NIHSS score ($p < 0.0001$). Patient comorbidities were not significantly different between outcome groups (**Supplementary Table S1**) and therefore were not included in the development of ML models.

Non-contrast CT imaging features associated with favorable outcomes included greater e-ASPECTS ($p = 0.002$), larger brain volume ($p = 0.043$), smaller cortical CSF volume ($p < 0.0001$), smaller lateral ventricle volume ($p < 0.0001$), smaller acute ischemic volume ($p = 0.047$), and non-acute ischemic volume ($p = 0.040$). Collateral circulation deficit volume on CTA was significantly lower in the favorable outcome group ($p = 0.001$; **Table 1**).

ML Model

Performance evaluation of each ML model following grid search optimization is presented in **Supplementary Table S2**. The selected features were as follows: baseline NIHSS, age, occlusion side, local M5 infarct volume, local lentiform infarct volume, brain volume, percentage of lateral ventricle volume, collateral vessel deficit volume, and the time interval from symptoms onset to admission.

We calculated accuracy and AUC for each of the feature groups on the training data set. For clinical data, the model with the highest AUC score was the XGBoost classifier (XGB) with an AUC of 81%. For imaging features, the best model was XGB at 79% AUC. For combined clinical and imaging features, the best model was also XGB with an AUC of 80%. Selected features yielded an AUC of 84% (**Figures 2A–D**).

The XGB model was selected for further optimization as it consistently achieved a high performance in the four feature groups and had the highest overall AUC scores. Using Bayesian hyperparameter tuning with a stratified cross-validation of 10-folds to refine the XGB model, the final performance metrics on the testing set of patients were AUC = 84%, accuracy = 77%, F1-score (mRS ≤ 2) = 67%, and F1-score (mRS > 2) = 82% for the selected features. The final results are shown in **Figures 2E,F**.

Following prediction modeling of mRS-90, a feature importance rank for the patient cohort was established by calculating SHAP values for XGBoost which revealed that

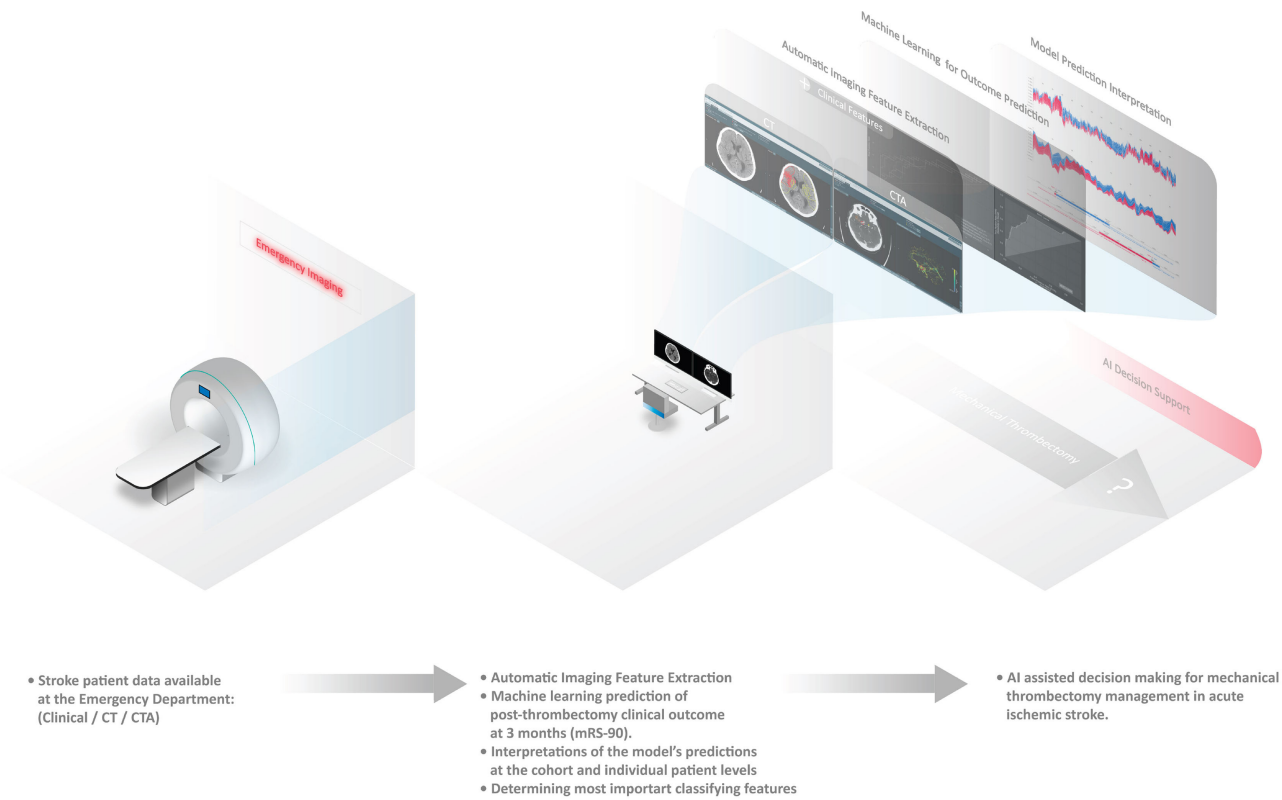


FIGURE 1 | Automated pipeline system for stroke functional outcome prediction at the emergency imaging providing artificial intelligence (AI) decision support for mechanical thrombectomy.

the top indicators of clinical outcome prediction for the model were by order of importance, i.e., age, baseline NIHSS, occlusion side, cortical CSF volume percentage, lateral ventricle volume percentage, e-ASPECTS, and circulation deficit volume (**Figure 3E**). The overall impact of each feature is represented by feature SHAP general values as shown in **Figure 3F** and model predictions were able to be reviewed and assessed regarding each predictor for each patient instance. The SHAP force plot allows for an interactive visualization of all the study populations clustered by their feature value similarity and ranging according to their specific model output (**Figures 3A,B**). Individual patient predictions can be extracted to visualize which features played a role in their classification and what their feature values were. Examples of predictions for a patient with poor outcomes and a patient with favorable outcomes are shown in **Figures 3C,D**, respectively.

DISCUSSION

In this study, we have developed and tested ML models to predict the 3-month functional outcome of patients with AIS and large vessel occlusion treated with mechanical thrombectomy using only clinical and imaging features available in the emergency department. Employing very simple baseline clinical information and automatically extracting quantitative imaging features from the baseline CT and CTA, our final model achieved very good predictive accuracy. Some of the features incorporated into our predictive model had not been previously examined, such as radiological markers of brain atrophy (brain volume, cortical CSF volume, and ventricular volume). In addition, we presented the feasibility of building interpretable ML models for stroke outcome prediction. The reporting of our prediction model includes information on what features weighed more heavily on the prediction that the algorithm utilized to construct the model.

The high evaluation metrics results in our study could be attributed to the newly introduced quantifiable features from automated image post-processing technology and the use of Bayesian hyperparameter tuning. Although all the included features contributed to the model performance, the most important features for the final model outcome prediction were as follows: age, baseline NIHSS, occlusion side, degree of brain atrophy (primarily represented by cortical CSF volume and lateral ventricle volume), early ischemic core (primarily represented by e-ASPECTS), and circulation deficit volume on CTA. This demonstrates the opportunity for multiple automatic imaging biomarkers extractable from routinely acquired imaging modalities (CT and CTA) to improve the precision of patient profiling for AIS management.

The complexity of ML models leads to challenges in defining the reasoning behind their predictions, thus potentially

TABLE 1 | Statistical feature comparison between the two outcome groups.

Features	mRS-90 ≤ 2 (n = 101)	mRS-90 > 2(n = 192)	P-value
Clinical features			
Age, median (IQR)	63 (51–74)	75 (63–84)	<0.0001
Sex			0.016
Female, *n* (%)	39 (39%)	104 (54.2%)	
Male, *n* (%)	62 (61%)	88 (45.8%)	
NIHSS score, median (IQR)	13 (7–18)	18 (13–22)	<0.0001
Time to admission, median (IQR)	107 (68–186)	135 (68–302)	0.107
Imaging features			
Occlusion side			0.017
Right, *n* (%)	57 (56%)	79 (41%)	
Left, *n* (%)	44 (44%)	113 (59%)	
Occlusion location			0.280
ICA Terminus, *n* (%)	25 (25%)	55 (29%)	
M1, *n* (%)	50 (49%)	99 (51%)	
M2, *n* (%)	25 (25%)	34 (18%)	
M3, *n* (%)	1 (1%)	4 (2%)	
e-ASPECTS, median (IQR)	9 (8–10)	9 (7–10)	0.002
Acute ischemic Volume (mL), median (IQR)	9.14 (5–20)	12.52 (5–28)	0.047
Non-acute ischemic volume (mL), median (IQR)	0.39 (0–0)	0.50 (0–1)	0.040
Local acute ischemic volume			
M1 (mL), median (range)	0.0 (0.0–9.5)	0.0 (0.0–13.7)	0.006
M2 (mL), median (range)	0.3 (0.0–8.8)	0.8 (0.0–18.8)	0.005
M3 (mL), median (range)	0.0 (0.0–12.5)	0.4 (0.0–20.2)	0.002
M4 (mL), median (range)	0.0 (0.0–7.5)	0.0 (0.0–11.7)	0.004
M5 (mL), median (range)	0.8 (0.0–17.9)	1.3 (0.0–32.0)	0.034
M6 (mL), median (range)	0.1 (0.0–23.0)	0.8 (0.0–25.3)	0.012
Caudate (mL), median (range)	0.0 (0.0–2.6)	0.0 (0.0–2.6)	0.540
Insula (mL), median (range)	0.0 (0.0–8.0)	4.7 (0.0–8.0)	0.201
Internal capsule (mL), median (range)	0.0 (0.0–4.7)	0.0 (0.0–4.8)	0.744
Lentiform (mL), median (range)	2.3 (0.0–5.8)	2.6 (0.0–5.8)	0.679
Brain volume (L), mean (±SD)	1.30 (±0.16)	1.26 (±0.15)	0.043
Cortical CSF volume (%), median (IQR)	6.16 (4–9)	8.7 (6–10)	<0.0001
Lateral ventricle volume (%), median (IQR)	2.4 (1–3)	3.4 (2–5)	<0.0001
Circulation deficit volume, median (IQR)	15.8 (1–36)	30.42 (6–54)	0.001
CTA CS score, median (IQR)	3.0 (2–3)	2.0 (1–3)	<0.001

hampering clinical adoption. In this study, the SHAP methodology provided explicability to the ML model at the cohort level and for each individual patient prediction with user-friendly visualization tools for demonstration purposes. These types of approaches have the potential for resolving the "trust barrier" between clinicians and AI algorithms and could help to increase clinical engagement with ML as future practice tools.

Our work results are consistent with and validate findings from previous studies, which have evaluated ML models for outcome prediction after AIS. Jiang et al. illustrated that ML applied to clinical and advanced imaging features had superior performance in binary mRS-90 prediction when compared to the Stroke Prognostication using Age and NIHSS (SPAN-100) scale (12). They reported the best model AUC of 80% using the 6 best performing features that include CT perfusion features (baseline NIHSS, age, glucose at admission, ischemic core volume on CT perfusion, penumbra volume on CT perfusion, and CTA-clot burden score) (12). Brugnara et al. reported a model for predicting mRS-90 after endovascular treatment for AIS that achieved an AUC of 74% using just baseline clinical and radiological features (13). The incorporation of features from CT perfusion did not improve the predictive performance of their model, but the inclusion of angiographic and post-interventional features significantly improved the predictive performance with an AUC of 85%. The most important parameters for their mRS-90 prediction were NIHSS after 24 h, pre-morbid mRS, and volume of final infarction volume on post-interventional CT (13).

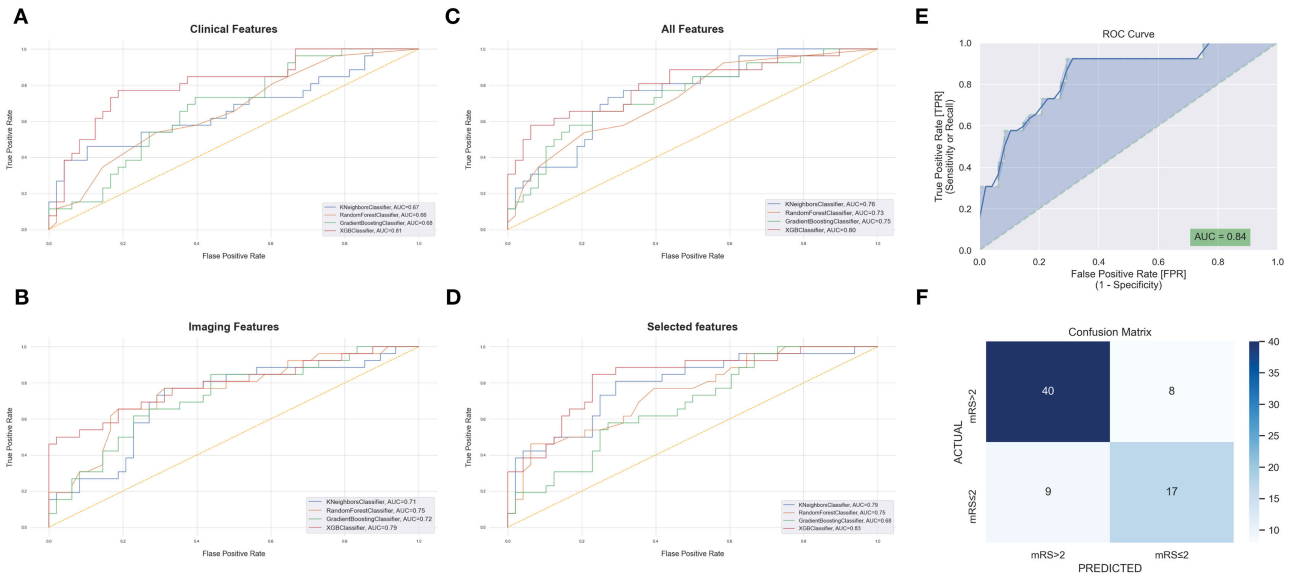


FIGURE 2 | Receiver operating characteristic curves (ROCs) with areas under curves for modified Rankin score at 90 days (mRS-90) prediction after grid-search optimization using baseline clinical features **(A)**, imaging features **(B)**, all features **(C)**, and selected features **(D)**, the orange-dashed line represents random guessing with an area under the receiver operating characteristic curve (AUC) of 0.5. The AUC **(E)** and confusion matrix **(F)** of the best performing model following Bayesian hyperparameter tuning using the selected features.

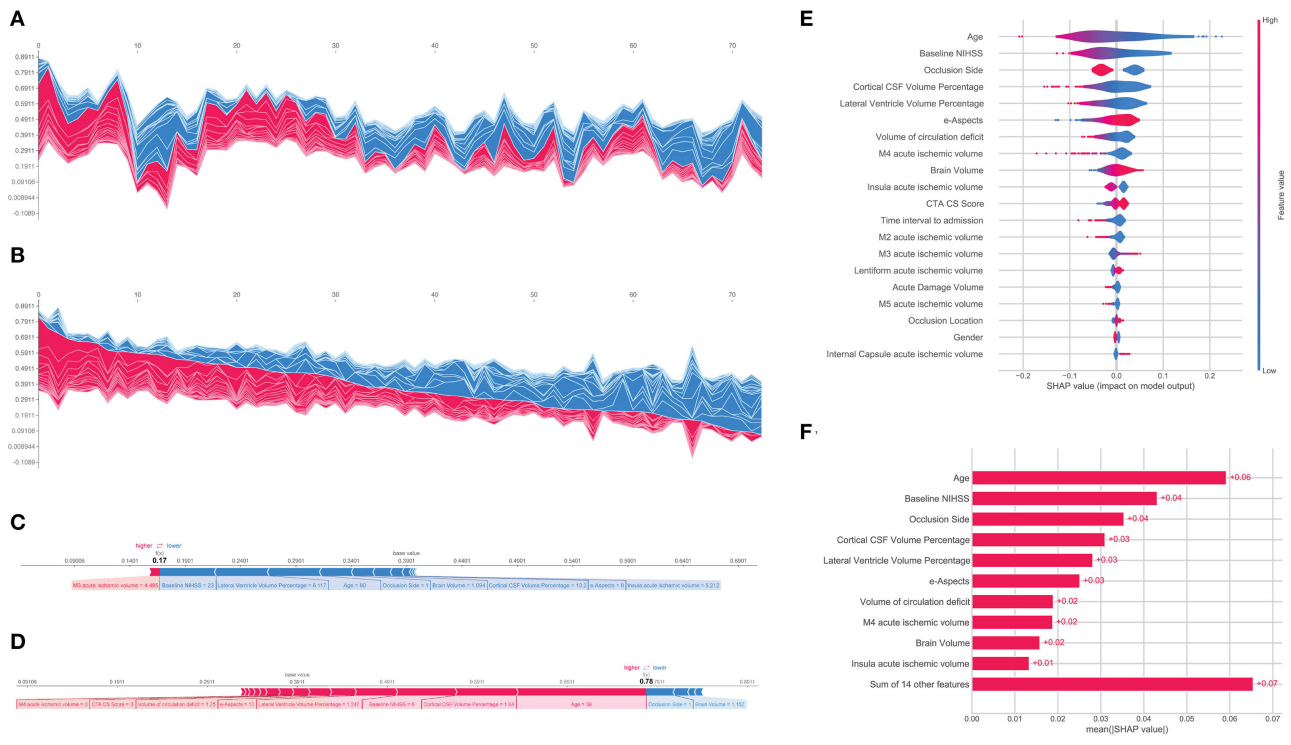


FIGURE 3 | Shapley Additive Explanation (SHAP) force plot of the testing set with the vertical axis representing model outcome and the horizontal axis representing the testing population sample ordered by feature similarity **(A)** and by model output **(B)**. Examples of the model output for an individual patient with the determining feature values that influenced the classification decision from the poor outcome group **(C)** and the favorable outcome group **(D)**. SHAP summary plot showing the distribution of each patient feature and how it affects the model outcome through its SHAP value **(E)**. Absolute mean SHAP values for the global effect of every feature effect on the model output **(F)**.

During model development, we experimented with incorporating features related to the endovascular intervention and post-interventional clinical and radiological features, such as TICI score, which as expected did increase the performance of the predictive models. However, the goal of the study was not just to merely develop a prognostic tool or achieve the highest prediction metrics possible but also to explore how the potential of ML models for decision supports in the setting of initial screening at the emergency department prior to the intervention and to identify the patients who would highly benefit if a mechanical thrombectomy procedure was to be performed. For that purpose, we intentionally chose to restrict the features to only those readily available at first scan and evaluation upon urgent patient arrival and excluded the post-interventional features from our final analysis. This interpretable approach could have promising applications and provide helpful service for stratifying patients with large vessel occlusion stroke prior to the endovascular procedure, leading to enhanced acute management decision-making.

Limitations

Limitations to our study may relate to the population size, which is although relatively large as a two-center study, similar to most ML studies, it could benefit from a larger cohort size for ML purposes. Datasets with diverse origins and a higher number of participants are warranted to further validate the robustness of the models for future generalizability on independent cohorts. In addition, with the absence of consistent information on pre-morbid functional status, we have not included this variable. Future planned steps exist for validating these tools prospectively and on larger multi-center datasets for further optimization of this approach.

CONCLUSION

Automated approaches could help to streamline and inform the decision-making process prior to thrombectomy in AIS at the emergency department. Our study highlights the value and accuracy of ML approaches integrating basic clinical information and automated imaging features in the pre-interventional prediction of functional outcomes 3 months from mechanical thrombectomy and the role of AI in both extracting useful information from routine imaging and individualizing prognostication and management decision-support systems in AIS. Progress made in ML interpretability is paving the way for more transparent modeling, which is becoming essential in the medical realm and for identifying important new predictors of stroke outcome.

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 Mayo Clinic Institutional Review Board (IRB). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

WB and DK contributed to conception and design of the study. GH and OJ provided software support for image post-processing. MJ performed the machine learning analysis, interpretation, visualization, and manuscript draft redaction. WB, DK, GH, AR, and TH supervised and contributed to manuscript revision, editing, and approval of the submitted version. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2022.884693/full#supplementary-material

REFERENCES

1. 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
2. Lin Y, Schulze V, Brockmeyer M, Parco C, Karathanos A, Heinen Y, et al. Endovascular thrombectomy as a means to improve survival in acute ischemic stroke: a meta-analysis. *JAMA Neurol.* (2019) 76:850–4. doi: 10.1001/jamaneurol.2019.0525
3. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* (2019) 50:e344–418. doi: 10.1161/STR.0000000000000211
4. 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. *New Engl J Med.* (2018) 378:708–18. doi: 10.1056/NEJMoa1713973
5. Campbell BCV, Donnan GA, Lees KR, Hacke W, Khatri P, Hill MD, et al. Endovascular stent thrombectomy: The new standard of care for large vessel ischaemic stroke. *Lancet Neurol.* (2015) 14:846–54. doi: 10.1016/S1474-4422(15)00140-4
6. 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
7. Drozdowska BA, Singh S, Quinn TJ. Thinking about the future: a review of prognostic scales used in acute stroke. *Front Neurol.* (2019) 10:e00274. doi: 10.3389/fneur.2019.00274
8. Raza SA, Rangaraju S. A review of pre-intervention prognostic scores for early prognostication and patient selection in endovascular management of large vessel occlusion stroke. *Interv Neurol.* (2018) 7:171–81. doi: 10.1159/000486539

9. Rangaraju S, Aghaebrahim A, Streib C, Sun C-H, Ribo M, Muchada M, et al. Pittsburgh response to endovascular therapy (pre) score: optimizing patient selection for endovascular therapy for large vessel occlusion strokes. *J Neurointerv Surg.* (2015) 7:783–8. doi: 10.1136/neurintsurg-2014-011351
10. Obermeyer Z, Emanuel EJ. Predicting the future — big data, machine learning, and clinical medicine. *New Engl J Med.* (2016) 375:1216–9. doi: 10.1056/NEJMp1606181
11. 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
12. Jiang B, Zhu G, Xie Y, Heit JJ, Chen H, Li Y, et al. Prediction of clinical outcome in patients with large-vessel acute ischemic stroke: Performance of machine learning versus span-100. *AJNR Am J Neuroradiol.* (2021) 42:240–6. doi: 10.3174/ajnr.A6918
13. Brugnara G, Neuberger U, Mahmutoglu MA, Foltyn M, Herweh C, Nagel S, et al. Multimodal predictive modeling of endovascular treatment outcome for acute ischemic stroke using machine-learning. *Stroke.* (2020) 51:3541–51. doi: 10.1161/STROKEAHA.120.030287
14. Brinjikji W, Abbasi M, Arnold C, Benson JC, Braksick SA, Campeau N, et al. e-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score. *Intervent Neuroradiol.* 27:781–7. doi: 10.1177/15910199211011861
15. Bouslama M, Ravindran K, Harston G, Rodrigues GM, Pisani L, Haussen DC, et al. Noncontrast computed tomography e-stroke infarct volume is similar to rapid computed tomography perfusion in estimating postreperfusion infarct volumes. *Stroke.* (2021) 52:634–41. doi: 10.1161/STROKEAHA.120.031651
16. Nagel S, Joly O, Pfaff J, Papanagiotou P, Fassbender K, Reith W, et al. e-ASPECTS derived acute ischemic volumes on non-contrast-enhanced computed tomography images. *Int J Stroke.* (2020) 15:995–1001. doi: 10.1177/1747493019879661
17. Herweh C, Ringleb PA, Rauch G, Gerry S, Behrens L, Möhlenbruch M, et al. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing ct scans of acute ischemic stroke patients. *Int J Stroke.* (2016) 11:438–45. doi: 10.1177/1747493016632244
18. Seker F, Pfaff JAR, Mokli Y, Berberich A, Namias R, Gerry S, et al. Diagnostic accuracy of automated occlusion detection in ct angiography using e-CTA. *Int J Stroke.* 17:77–82. doi: 10.1177/1747493021992592
19. Grunwald IQ, Kulikovski J, Reith W, Gerry S, Namias R, Politi M, et al. Collateral automation for triage in stroke: Evaluating automated scoring of collaterals in acute stroke on computed tomography scans. *Cerebrovasc Dis.* (2019) 47:217–22. doi: 10.1159/000500076
20. Tan IY, Demchuk AM, Hopyan J, Zhang L, Gladstone D, Wong K, et al. Ct angiography clot burden score and collateral score: Correlation with clinical and radiologic outcomes in acute middle cerebral artery infarct. *AJNR Am J Neuroradiol.* (2009) 30:525–31. doi: 10.3174/ajnr.A1408
21. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, editos. *Advances in Neural Information Processing Systems 30*. Long Beach, CA: Curran Associates, Inc. (2017). p. 4765–74.
22. 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
23. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng.* (2018) 2:749–60. doi: 10.1038/s41551-018-0304-0

Conflict of Interest: GH and OJ receive salary support and share options from Brainomix 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.

frontiers | Frontiers in Neurology

ORIGINAL RESEARCH
published: 19 July 2022
doi: 10.3389/fneur.2022.932933

Protocol and Preliminary Results of the Establishment of Intracranial Aneurysm Database for Artificial Intelligence Application Based on CTA Images

Wei You[1,2†], Yong Sun[3†], Junqiang Feng[1,2], Zhiliang Wang[4], Lin Li[5], Xiheng Chen[1,2], Jian Lv[1,2], Yudi Tang[1], Dingwei Deng[1], Dachao Wei[1], Siming Gui[1], Xinke Liu[1,2], Peng Liu[1,2], Hengwei Jin[1,2], Huijian Ge[1,2] and Yanling Zhang[6]*

[1] Department of Interventional Neuroradiology, Beijing Neurosurgical Institute and Beijing Tiantan Hospital, Capital Medical University, Beijing, China, [2] Department of Neurointerventional Engineering and Technology, Beijing Engineering Research Center, Beijing, China, [3] Department of Neurosurgery, The First People's Hospital of Lianyungang and The First Affiliated Hospital of Kangda College, Nanjing Medical University, Lianyungang, China, [4] Department of Neurosurgery, The People's Hospital of Nanpi Country, Cangzhou, China, [5] Department of Emergency, The Fourth Hospital of Hebei Medical University, Shijiazhuang, China, [6] Department of Roradiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Baofeng Gao,
Beijing Institute of Technology, China

Reviewed by:
Ji Wenjun,
Yulin No.2 People's Hospital, China
Wei Ni,
Fudan University, China

***Correspondence:**
Yanling Zhang
zyl5545@126.com

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

Specialty section:
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

Received: 30 April 2022
Accepted: 20 June 2022
Published: 19 July 2022

Citation:
You W, Sun Y, Feng J, Wang Z, Li L, Chen X, Lv J, Tang Y, Deng D, Wei D, Gui S, Liu X, Liu P, Jin H, Ge H and Zhang Y (2022) Protocol and Preliminary Results of the Establishment of Intracranial Aneurysm Database for Artificial Intelligence Application Based on CTA Images. Front. Neurol. 13:932933. doi: 10.3389/fneur.2022.932933

Background and Purpose: Unruptured intracranial aneurysms (UIAs) are increasingly being detected in clinical practice. Artificial intelligence (AI) has been increasingly used to assist diagnostic techniques and shows encouraging prospects. In this study, we reported the protocol and preliminary results of the establishment of an intracranial aneurysm database for AI application based on computed tomography angiography (CTA) images.

Methods: Through a review of picture archiving and communication systems, we collected CTA images of patients with aneurysms between January 2010 and March 2021. The radiologists performed manual segmentation of all diagnosed aneurysms on subtraction CTA as the basis for automatic aneurysm segmentation. Then, AI will be applied to two stages of aneurysm treatment, namely, automatic aneurysm detection and segmentation model based on the CTA image and the aneurysm risk prediction model.

Results: Three medical centers have been included in this study so far. A total of 3,190 cases of CTA examinations with 4,124 aneurysms were included in the database. All identified aneurysms from CTA images that enrolled in this study were manually segmented on subtraction CTA by six readers. We developed a structure of 3D-Unet for aneurysm detection and segmentation in CTA images. The algorithm was developed and tested using a total of 2,272 head CTAs with 2,938 intracranial aneurysms. The recall and false positives per case (FP/case) of this model for detecting aneurysms were 0.964 and 2.01, and the Dice values for aneurysm segmentation were 0.783.

Conclusion: This study introduces the protocol and preliminary results of the establishment of the intracranial aneurysm database for AI applications based on CTA

images. The establishment of a multicenter database based on CTA images of intracranial aneurysms is the basis for the application of AI in the diagnosis and treatment of aneurysms. In addition to segmentation, AI should have great potential for aneurysm treatment and management in the future.

Keywords: intracranial aneurysm, database, artificial intelligence, computed tomographic angiography, deep learning

BACKGROUND

Saccular unruptured intracranial aneurysms (UIAs) are pathological artery dilation that occurs in major cerebral arteries branch and affects 3–5% of the adult population (1). Approximately 20–30% of patients with intracranial aneurysms harbor more than one aneurysm (2). With the improvement in the quality of intracranial imaging techniques over the past 20 years and the more widespread application of magnetic resonance imaging (MRI) and computed tomography (CT) as diagnostic tools (3), UIAs are increasingly being detected in clinical practice. A recent cross-sectional study showed that the prevalence of UIA is as high as 7% among individuals aged 35–75 years in China (4). Subarachnoid hemorrhage caused by aneurysm rupture has serious consequences, with the mortality rate of early hemorrhage being 40%, and the rate of rebleeding being as high as 60–70% (5). Surgical clipping or interventional therapy for aneurysms is associated with an inherent risk of invasions, with a 4.3–4.6% incidence of post-operative morbidity and a 10–24.6% incidence of new neurological deficits following treatment (6). Clinicians are increasingly faced with the dilemma of choosing appropriate clinical management, either prophylactic treatment (endovascular or aneurysm clipping) with an inherent risk of complications or conservative treatment that leaves patients at risk of aneurysm rupture. Establishing a reliable model to determine the stability of UIAs is important for therapeutic decisions in unruptured aneurysms.

IMAGING DIAGNOSTIC TOOLS FOR INTRACRANIAL ANEURYSMS

The detection and risk assessment of intracranial aneurysms is critical due to their low rupture rate and high rates of disability and mortality after rupture. IAs are most often detected incidentally after the rupture of aneurysms or during the evaluation of systemic symptoms, such as headache, neural paralysis, and ischemic cerebrovascular disease. At present, computed tomography angiography (CTA), magnetic resonance angiography (MRA), and digital subtraction angiography (DSA) are the main imaging diagnostic tools for intracranial aneurysms. Each tool has advantages and disadvantages, and different individual practitioners use them variably at various stages in the evaluation of an aneurysm.

As the "gold standard" for aneurysm diagnosis, DSA has shown greater sensitivity, especially in aneurysms smaller than 3 mm (7, 8). However, it is difficult to become a large-scale disease screening method due to its inherent invasiveness. MRA for aneurysm imaging uses contrast methods or time-of-flight (TOF) sequences. In general, MRA has been reported to have a detection sensitivity ranging from 74 to 98% (9). However, a study showed that the sensitivity of MRA can be significantly affected by the aneurysm size. Therefore, magnetic resonance was suggested as a primary method of screening for UIAs and can be very useful for aneurysms larger than 3 mm. Possible complications for DSA include contrast agent allergy events, IA rupture, brain infarction, and arterial injury (10). With the development of multidetector scanners, CTA is frequently added to assist in the diagnosis of the aneurysm. In general, the sensitivity, specificity, and accuracy of aneurysm detection are very high compared with DSA with the 3D rotational acquisition, with 1 report indicating values of 96.3, 100, and 94.6%, respectively (11). Even for smaller aneurysms (typically for those <3 mm), there was a comparable sensitivity, specificity, and accuracy (81.8, 100, and 93.3%, respectively). Therefore, with its high sensitivity and specificity, CTA can be considered an initial diagnostic test for aneurysm detection and screening (12).

Management of UIAs and Predictors of Rupture

After a UIA is detected, several factors must be considered to determine the appropriate management approach. Despite the continuous improvement of surgical clipping and endovascular treatment techniques, the therapeutic effect of the intracranial aneurysm has been gradually improved, but there are still certain complications and morbidity mortality in various treatment methods. There are no randomized clinical trial data that define the optimum management of a UIA. The available natural history studies provide both retrospective and prospective data. The risk of aneurysm rupture without any intervention should be compared with the risk of surgical clipping or endovascular treatment. The need for preventive treatment of unruptured aneurysms is controversial, especially considering the low rupture rate of intracranial aneurysms, which is only 0.25% (12). Therefore, determining its stability is important to make therapeutic decisions for unruptured aneurysms (13, 14). The most reasonable treatment strategy for intracranial aneurysms may be to screen patients with a high risk of rupture and carry out active intervention treatment and to carry out conservative and follow-up treatment for patients with a low risk of rupture.

Many studies have identified important clinical factors for aneurysm rupture. Size is proposed as the most important factor for predicting aneurysm stability in previous studies. Up to 85–90% of ruptured intracranial aneurysms are <10 mm in diameter (15–17). According to the International Study of Unruptured

intracranial aneurysms, aneurysms of <7 mm in the anterior and posterior circulation have a cumulative 5-year risk of rupture of 0 and 2.5%, respectively (18). However, a published lifelong follow-up study that smoking and female gender were more severe prognostic factors than aneurysm size (19). Old age was also determined as a risk factor for aneurysm rupture in a prospective study (20). As growing aneurysms are at high risk for rupture compared with stable size and stable shape (21), discriminating aneurysm stability is meaningful. In addition to size and patient's baseline information, the morphology of the aneurysm is closely related to the stability of the aneurysm (13, 14). Some studies have found that various morphological features of aneurysms, such as size ration, flow angle, height/width ratio, aspect ratio, and deviated angle, are associated with their rupture. However, these parameters are mainly measured in a two-dimensional projection and may differ among different angle projections or measurers.

It is well-known that three-dimensional images contain primitive and comprehensive morphological information, and radiomics and machine learning studies based on these images have attempted to identify important morphological features associated with aneurysm stability. In a study containing 719 aneurysms, the researchers used radiomics to find that SphericalDisproportion, Maximum 2D diameter slice, and surface area were the most important morphological feature to predict aneurysm stability (22). However, the research on these complex three-dimensional parameters is still incomplete, and there are obvious problems, such as insufficient scientific research data and inconsistent experimental results. Their values of aneurysm rupture risk assessment and prediction need to be further verified (14, 23).

Therefore, it is urgent to explore the risk factors closely related to intracranial aneurysm rupture that is convenient for quantitative analysis and to establish an accurate and effective risk assessment system for intracranial aneurysm rupture on this basis, providing a reliable basis for treatment decisions, which has great clinical and social value.

Electronic Medical Data

At present, almost all tertiary medical centers in China use electronic medical record systems and picture archiving and communication system (PACS). Digitization of the medical records and the storage of image data in the "DICOM" format provides an opportunity to utilize medical data for standardization and analysis in unique ways. In addition, modern tools used for extraction from electronic medical records and image data from PACS can be used to develop a database that can be used for retrospective studies and registry studies and can even be used for establishing stratified disease models. Digitization of medical data makes it possible to analyze imaging data of a large number of patients based on medical records. Although there are a large number of patients with intracranial aneurysms and a large amount of relevant medical data in medical centers in China, there are currently data barriers between hospitals, making data sharing difficult. The establishment of a multicenter database is the foundation for high-quality, large-scale disease research in the future.

Artificial Intelligence

The promise of artificial intelligence (AI) in medicine and healthcare is widely discussed at present, just as it was when the concept was first proposed in the late 1960s. AI focuses exclusively on the use of algorithms and the software that implements them to approach human cognition abilities to analyze complex observations and data. AI works by modeling and extracting information from various data sources and then processing it, with the goal of being able to provide a well-defined, ideally understandable, and interpretable output to the practicing medical professional. It encompasses different analytical methods, such as machine learning (ML), natural language processing, and computer or machine vision (24). ML, represented by deep learning, represents the most successful branch of AI, focusing on the development of programs with the ability to learn from data (25).

It has recently been used to aid in diagnostic techniques, and several studies have attempted to apply ML methods to neuroimaging data to assist in stroke diagnosis (26). The relevant published literature has focused on several disease types, namely, cancer, neurological diseases, and cardiovascular diseases (26). The application of AI relies on massive amounts of medical data being collected and stored in the form of electronic medical records, especially including rich medical imaging information. A major advantage of deep learning is the convolutional neural network (CNN), which analyzes pixel-level information in images and is able to interpret the orientation of pixels, which allows them to identify lines, curves, and even objects in images. A study shows that the applications of AI in precision oncology are due in large part to its remarkable ability to classify imaging data across different clinical domains (25). Data mining has great significance for identifying new diagnostic markers that can precisely diagnose and classify diseases, such as intracranial aneurysms, based on information about them.

METHODS AND DESIGN

Building an IA Database Based on the CTA Image for AI

We aim to apply the full advantage of AI to the diagnosis and treatment of aneurysms, including automatic diagnosis of aneurysms, automatic segmentation, and risk assessment of aneurysms. Therefore, we will establish a multicenter database based on CTA images of intracranial aneurysms, which is the basic work of automatic aneurysm diagnosis, segmentation, and aneurysm rupture risk prediction based on AI technology. To construct the database, previous retrospective data and prospective data enrolled during the registration of studies will be included.

The reason why we are willing to explore this issue is due to the urgent clinical needs: (1) compared with the limited physician resources, there are a large number of aneurysm images to be diagnosed; (2) more reliable aneurysm risk assessment methods are needed to develop reasonable treatment strategies. Also at the technological frontier, we are seeing clear reasons, including data availability and expert opinion, that this exploration is feasible.

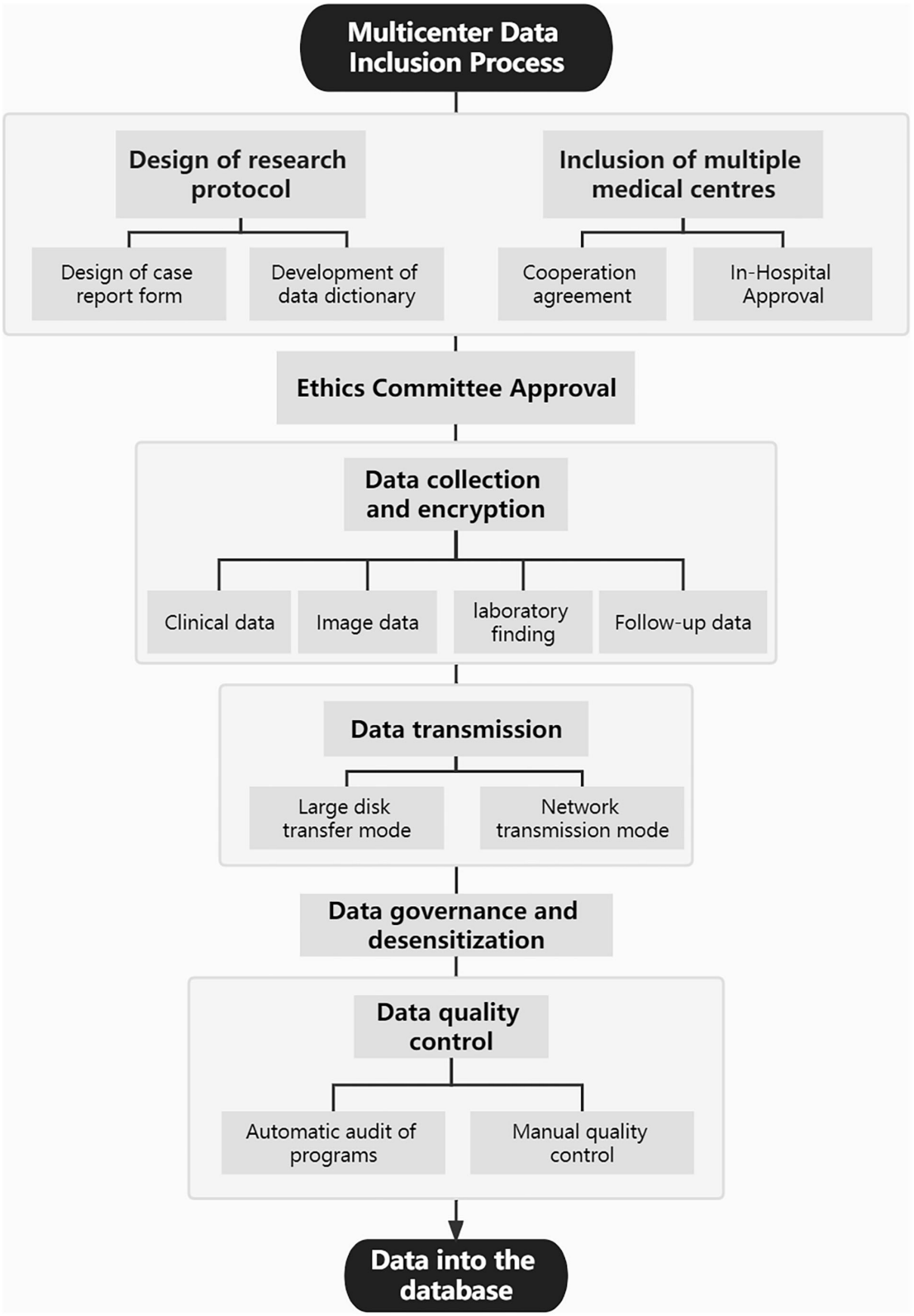


FIGURE 1 | The process of multicenter data inclusion.

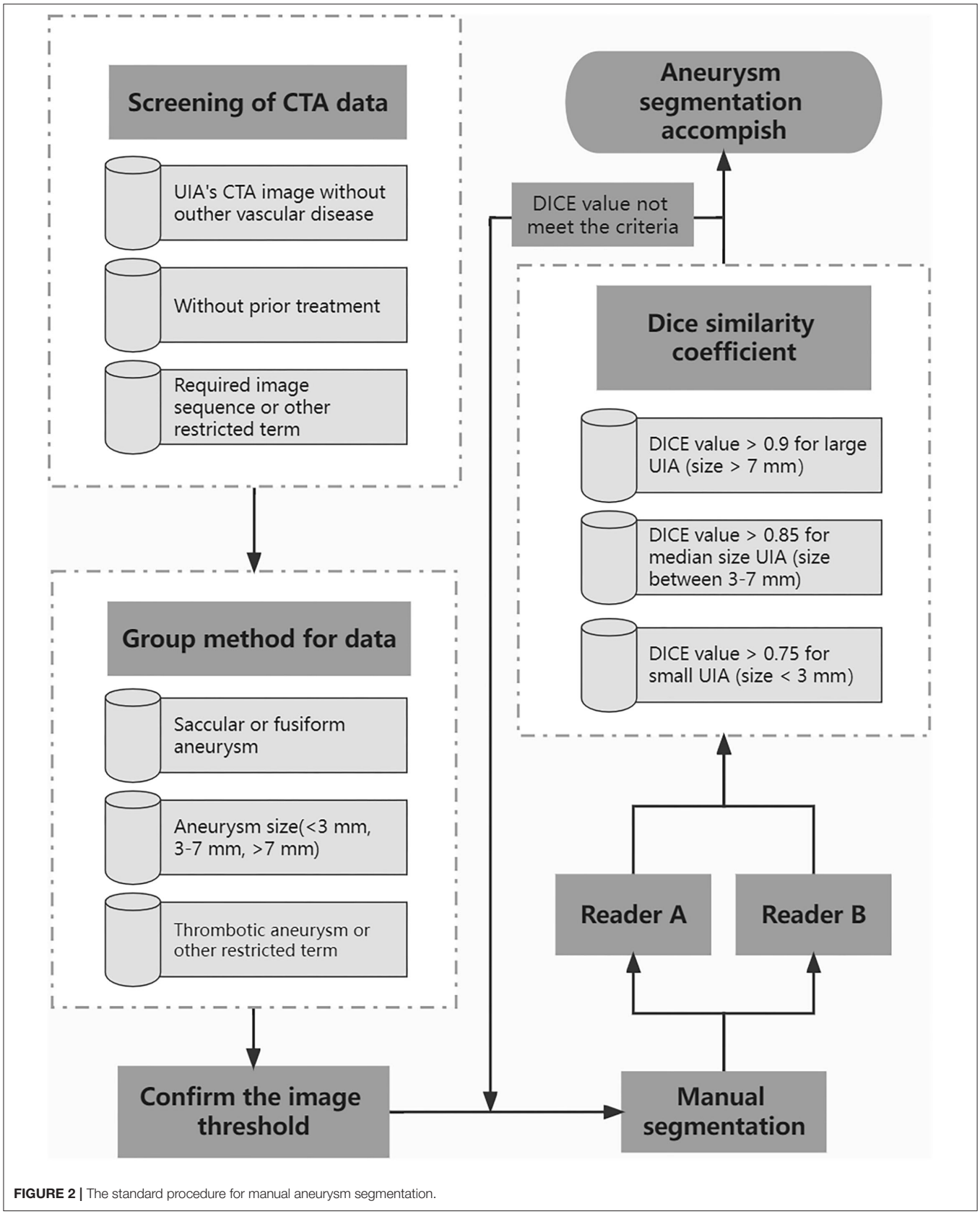


FIGURE 2 | The standard procedure for manual aneurysm segmentation.

This study was approved by the medical ethics committee of Tiantan Hospital, Beijing, China. Through a review of the picture archiving and communication system in Tiantan hospital, we collected CTA images of all patients with aneurysms treated in our hospital between January 2010 and March 2021. In addition, CTA data of patients with aneurysms from two other Chinese medical centers were also collected. Furthermore, on the basis of all enrolled patients with aneurysms with CTA images, the corresponding clinical data were collected through the electronic medical record systems, including basic clinical characteristics, relevant data concerning personal and family history, relevant laboratory findings, and follow-up information. **Figure 1** shows the multicenter data inclusion process. Data continue to be added through observational registries (registration number: ChiCTR2100054564). Patients' treatment was not affected by their participation in this study.

Recruitment of Participants

Collection Criteria

At least one UIA that has not been treated and has been confirmed by imaging: CTA, MRA, and/or DSA.

Complete CTA image sequence and clinical data.

Patients or relatives agreed to participate in this study.

Exclusion Criteria

Computed tomography angiography with arteriovenous malformation, arteriovenous fistula, or Moyamoya disease.

Post-traumatic or infectious pseudoaneurysm.

Previous surgical clipping or endovascular treatment for the aneurysm.

Non-diagnostic image quality with severe artifacts as judged by an attending neuroradiologist.

Incomplete or missing CTA imaging data and clinical data.

Manual Segmentation of Aneurysm for Model Training

A standard procedure for manual aneurysm segmentation (**Figure 2**) was developed through consultation between doctors and technical engineers, and the test of the standard was completed through a small amount of data. All operators were trained in a standardized manner prior to formal manual segmentation of aneurysms. Eligible CTA data will be grouped on each slice on the basis of aneurysm characteristics. Identified aneurysms were manually segmented on subtraction CTA by radiologists using 3D Slicer (version 4.10.1). Each data will be randomly assigned to two readers at the

TABLE 1 | The overview of the characteristics of the database.

Characteristic	Total
No. of patients, *n*	3,190
No. of IAs, *n*	4,124
Male, *n* (%)	1,219 (38.2)
Age, years	56.7 ± 11.0
Male, years	55.8 ± 11.5
Female, years	57.2 ± 10.6
Aneurysm size, mm	2.3 ± 0.5
<3 mm, *n* (%)	1,540 (37.3)
3–7 mm, *n* (%)	2,035 (49.3)
>7 mm, *n* (%)	545 (13.2)
Aneurysm's location	
ICA, *n* (%)	2,384 (57.8)
ACA, *n* (%)	227 (5.5)
ACoA, *n* (%)	403 (9.8)
MCA, *n* (%)	696 (16.9)
PCA, *n* (%)	89 (2.2)
PCoA, *n* (%)	28 (0.7)
BA, *n* (%)	151 (3.7)
VA, *n* (%)	109 (2.6)
Other, *n* (%)	37 (0.9)
Scanners	
GE healthcare, *n* (%)	2,699 (84.6)
Philips scanners, *n* (%)	482 (15.1)
Other scanner, *n* (%)	9 (0.3)

ICA, internal carotid artery; ACA, anterior cerebral artery; ACoA, anterior communicating artery; MCA, middle cerebral artery; PCA, posterior cerebral artery; PCoA, posterior communicating artery; BA, basilar artery; VA, vertebral artery.

TABLE 2 | The overview of the characteristics of each dataset.

Characteristic	Training set	Validation set	Test set	Total
No. of patients, *n*	1,606	314	352	2,272
No. of IAs, *n*	2,078	414	446	2,938
CTA with single IA, *n*	1,233	237	283	1,753
CTA with multiple IAs, *n*	373	77	69	519
Female, *n* (%)	1,008 (62.8)	183 (58.3)	216 (61.4)	1,407 (61.9)
Age, years	56 ± 10	57 ± 12	55 ± 10	56 ± 11
Male	56 ± 11	57 ± 11	53 ± 10	56 ± 11
Female	57 ± 10	57 ± 12	56 ± 10	57 ± 10
Aneurysm size, mm	4.2 ± 3.4	3.6 ± 2.5	6.0 ± 4.6	4.4 ± 3.6
<3 mm, *n* (%)	880 (42.3)	209 (50.5)	94 (21.1)	1,183 (40.3)
3–7 mm, *n* (%)	982 (47.3)	174 (42)	242 (54.3)	1,398 (47.6)
>7 mm, *n* (%)	216 (10.4)	31 (7.5)	110 (24.7)	357 (12.2)
Aneurysm's location				
ICA, *n* (%)	1,344 (64.7)	258 (62.3)	171 (38.3)	1,773 (60.3)
ACA, *n* (%)	98 (4.7)	21 (5.1)	30 (6.7)	149 (5.1)
ACoA, *n* (%)	143 (6.9)	23 (5.6)	65 (14.6)	231 (7.9)
MCA, *n* (%)	299 (14.4)	65 (15.7)	136 (30.5)	500 (17)
PCA, *n* (%)	50 (2.4)	7 (1.7)	6 (1.3)	63 (2.1)
PCoA, *n* (%)	13 (0.6)	2 (0.5)	10 (2.2)	25 (0.9)
BA, *n* (%)	66 (3.2)	20 (4.8)	11 (2.5)	97 (3.3)
VA, *n* (%)	53 (2.6)	17 (4.1)	14 (3.1)	84 (2.9)
Other, *n* (%)	12 (0.6)	1 (0.2)	3 (0.7)	16 (0.5)
Scanners				
GE healthcare, *n* (%)	1,601 (99.7)	314 (100)	352 (100)	2,267 (99.8)
Philips scanners, *n* (%)	5 (0.3)	0 (0)	0 (0)	5 (0.2)

ICA, internal carotid artery; ACA, anterior cerebral artery; ACoA, anterior communicating artery; MCA, middle cerebral artery; PCA, posterior cerebral artery; PCoA, posterior communicating artery; BA, basilar artery; VA, vertebral artery.

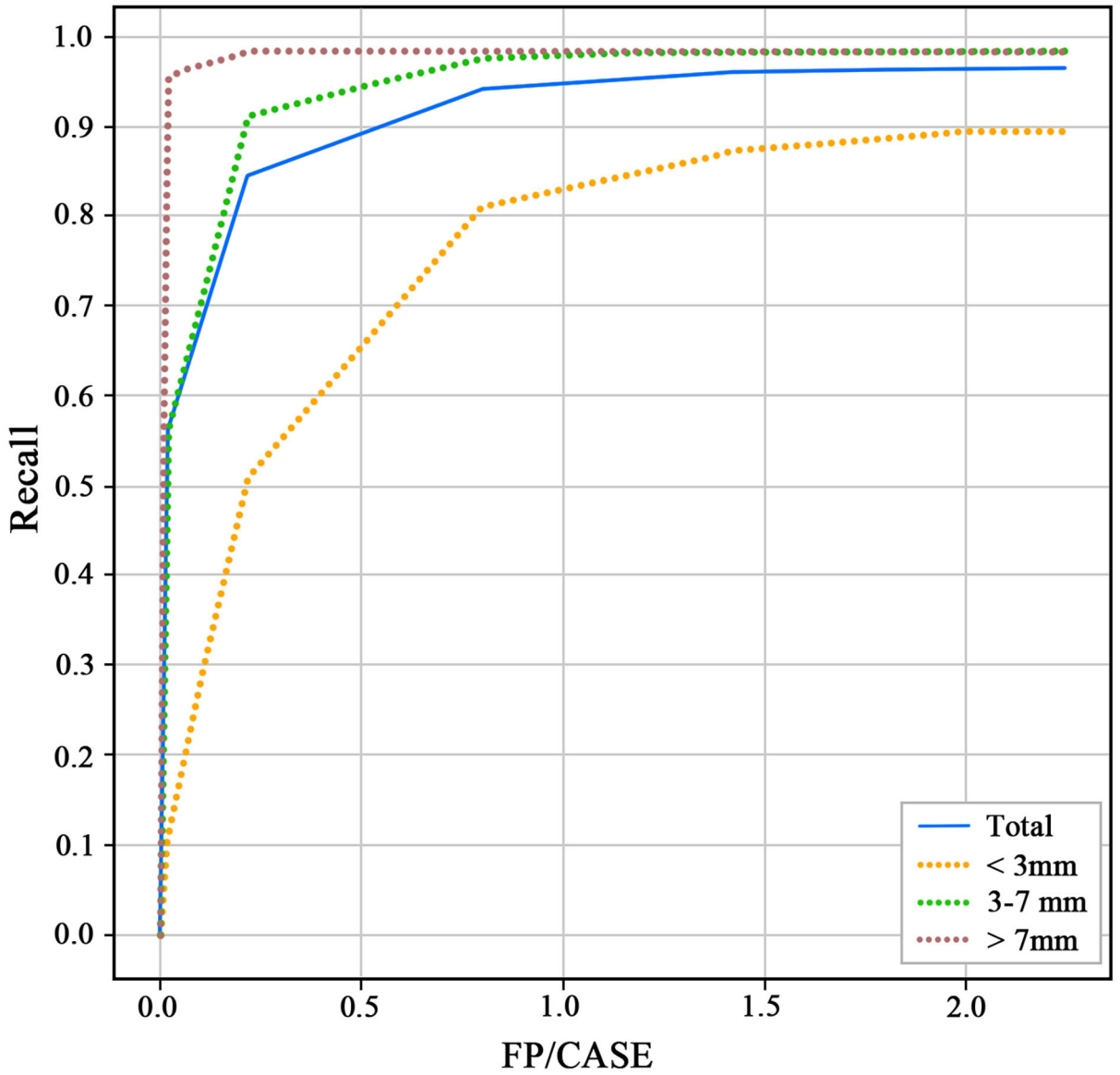


FIGURE 3 | Receiver operating characteristics (ROC) of different size aneurysms for the test dataset.

same time for segmentation using the same threshold. Dice similarity coefficient (DSC) was used to evaluate the annotation performance of two readers. For the aneurysm sizes >7, 3–7, and <3 mm, the DSC criteria for segmentation were >0.9, >0.85, and >0.75, respectively. If the DSC does not meet the criteria, it will be returned for re-annotation until meets the requirement.

Aneurysm Detection and Segmentation Model

All data included in this study were divided into training, validation, and test datasets in a ratio of 7:1.5:1.5. The training dataset was used to develop the 3D-Unet model, the validation dataset was used to validate the algorithm and adjust the model hyperparameters, and the test dataset was used to evaluate the generalization ability of the model. Manual aneurysm detection and segmentation were used as the reference standard to evaluate the performance of the final model.

For the study on algorithm development, we developed a structure of 3D-Unet for aneurysm segmentation in CTA images. The network takes CTA as input and outputs a probability mask, which contains the probability of whether each voxel belongs to the aneurysm. Each CTA image was cropped into patches with the same size of (32, 224, 224) for training the 3D-Unet. Hounsfield units of each patch image were normalized to (0, 1). Data augmentations methods, such as random crop and random zoom, were used for robust training of the network. The total loss

was the sum of the losses of the main output and module and minimized using the Adam Optimizer.

The development of automated detection models for aneurysms has the potential to reduce reading time and increase radiologists' performance in laboratory and clinical environments. The detection model may also benefit patients undergoing CTA (those with a headache or acute ischemic stroke) because it may reduce the likelihood of an incidental aneurysm being undetected. Additionally, the realization of automatic segmentation of aneurysms will establish a good foundation for the risk prediction model in the next stage. It can provide a large amount of learnable data for risk prediction models in learning aneurysm morphological features in a short time, breaking through the limitation of the high labor cost of manual segmentation. Furthermore, it can reduce observer variability and avoid bias between observers.

Risk Prediction Models for Aneurysms

The risk of IA can be divided into three stage-related factors: the risk of aneurysm development, the risk of growth or morphological changes, and the risk of rupture. Among them, we focused on the risk of growth and rupture following aneurysm development. We divided all aneurysm data into the stable group and unstable group based on morphological changes of aneurysms under multiple imaging follow-up examinations and aneurysm hemorrhage. Aneurysm data in the unstable group included ruptured and growing aneurysms, which are at high risk. The risk prediction of aneurysms is a complex problem, and the related predictive factors reported in the literature mainly include patient clinical factors, aneurysm morphological parameters, and hemodynamic parameters. Using statistical models to build a multidimensional model for UIA risk prediction that integrates clinical, image, and hemodynamic data is urgently needed in clinical practice. In this process, the deep involvement of AI is the key to the effective use of massive data and the establishment of efficient and accurate models. AI technology will be used for the automatic calculation of morphological parameters based on the automatic segmentation of aneurysms. AI can also use to automatically extract valuable clinical information from a large amount of chaotic clinical data. Additionally, it can realize the rapid and automatic extraction of the hemodynamic parameters of the aneurysm through a large amount of learning of the computational fluid dynamics (CFD) analysis data of the aneurysm. The availability of expert annotations makes this a classic supervised ML problem, although we will also consider exploring semi-supervised and unsupervised learning methods as we continue to add data.

Statistical Analysis

Normality assumptions were assessed using the Kolmogorov-Smirnoff test. Data are presented as frequencies for categorical variables and means and ranges for continuous variables. Pearson's chi-square test or Fisher's exact test were used to evaluate the differences between ordinal and categorical variables. Independent sample tests were used to examine group differences of continuous variables. A two-sided *P*-value <0.05 was considered statistically significant. To evaluate the performance of the automatic segmentation model, results were expressed as Dice values. To evaluate the performance of automatic detection, the recall and false positives per case (FPs/case) of all aneurysms were evaluated in the test dataset. Multiple regression analysis will be used to calculate OR value for high-risk factors or significant parameters. A multidimensional model for UIA risk prediction will be used to assess the risk of aneurysm rupture and growth in the prospective cohort group and thus determine the model's accuracy and efficacy. Statistical analysis was performed using SPSS 22 (IBM, Chicago, IL, USA), and figures were generated using GraphPad Prism (GraphPad Software, San Diego, CA, USA).

RESULTS

Data Inclusion

Three medical centers, including Beijing Tiantan Hospital, have been included in this study so far. In a retrospective review of the picture archiving and communication systems at Beijing Tiantan Hospital from January 2010 to March 2021, CT data of a total of 7,855 patients with radiologically reported intracranial aneurysms were obtained. Among 2,896 CT without CTA, 1,671 CT was not from the head scan, 66 CTA was with arteriovenous malformation, fistula, and moyamoya disease, 230 CTA showed with previous treatment, and 11 CTA was with poor image quality with severe artifacts, thus excluded from this database. A total of 2,981 cases of CTA examinations with 3,872 aneurysms were included from Tiantan Hospital in the database. In the same way, a total of 209 CTA examinations with 252 aneurysms enrolled the database from the two other medical centers. In the end, a total of 3,190 CTA examinations with 4,124 aneurysms included in the database. The overview of the characteristics of the database is shown in **Table 1**. All identified aneurysms from CTA images that enrolled in this study were manually segmented on subtraction CTA by 6 readers, and the results met the criteria for segmentation. Additional medical centers and aneurysms data will continue to be added to the database in the future through the registry study.

Algorithm Development

At present, we have applied deep learning methods to the diagnosis and segmentation of aneurysms on CTA images. In this part, a total of 2,272 CTA examinations with 2,938 aneurysms were enrolled in this algorithm development. The training dataset contained 1,606 CTA examinations with 2,078 aneurysms, the validation dataset contained 314 CTA examinations with 414 aneurysms, and the test dataset contained 352 CTA examinations with 446 aneurysms. It should be noted that all cases from the test dataset underwent CTA examinations verified by DSA examination within 1 month. The characteristics of each dataset are shown in **Table 2**.

The 3D-Unet algorithm was developed with the training and validation datasets. After the training procedure, the hyperparameters with the best sensitivity on the validation dataset were chosen. For the test dataset, the recall and FP/case of the model to detect aneurysms were 0.964 and 2.01 (**Figure 3**). For the aneurysm sizes <3, 3–7, and >7 mm, the recall was 0.894,

0.983, and 0.982, respectively. The Dice similarity coefficient was used to evaluate the model performance of IA segmentation. For the test dataset, the Dice values are 0.783. For the aneurysm sizes <3, 3–7, and >7 mm, the Dice values are 0.635, 0.796, and 0.868, respectively.

DISCUSSION

Patient's clinical data and relevant image data house the most original and important data. With the improvement of the quality of intracranial imaging technology and the wide application of CT as a diagnostic tool, the clinical radiological examination for neurologic diagnoses is increasing, which requires human expertise in image interpretation. However, there is a relative shortage of experienced radiologists due to the increased demand for imaging diagnoses (27). As a result, there can be uncertainty and inevitable mistakes when making diagnoses and decisions. Large volumes of medical and imaging data are particularly suitable for the application of advanced computing technologies, especially AI and the related field of machine learning. With its high sensitivity and specificity and less invasiveness, CTA has been considered an initial diagnostic test for aneurysm screening and has generated a large amount of data in clinical practice, thus especially suitable for the development of an automatic diagnosis model of the aneurysm.

Our goal is to apply the full advantage of AI to the diagnosis and treatment of aneurysms. Therefore, the establishment of a multicenter database based on CTA images of intracranial aneurysms is the basis for the application of AI in the diagnosis and treatment of aneurysms. The reason why we are willing to explore this issue is due to the urgent clinical needs: (1) compared with the limited physician resources, there are a large number of aneurysm images to be diagnosed; (2) more reliable aneurysm risk assessment methods are needed to develop reasonable treatment strategies. Also at the technological frontier, we are seeing clear reasons, including data availability and expert opinion, that this exploration is feasible.

In addition to segmentation, AI should have great potential for aneurysm treatment and management in the future. Quantitative measurement of aneurysm morphological parameters and prediction of treatment risk and post-operative complications are potential image-based AI applications. In addition to the image data, this study also discusses the clinical baseline data and hemodynamic factors with the aim of making them more applicable to the clinical situation. Additionally, since the study only included Chinese individuals, the results will be more applicable to Chinese people, and the applicability to other populations is unknown.

CONCLUSION

A representative database comprising of multicenter based on CTA images of intracranial aneurysms is developed and presented in this work. The database developed comprises the results of segmentation performed by radiologists for each aneurysm, corresponding to the anatomical location and morphological characteristics of the aneurysm. Based on the database, the 3D-Unet algorithm was developed with the recall and FP/case of 0.964 and 2.01 to detect aneurysms and with the Dice values of 0.783 for aneurysm segmentation. In addition to automatic segmentation, AI should have great potential for future aneurysm treatment and management, such as aneurysm growth and rupture risk prediction.

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 Institutional Review Board of Beijing Tiantan Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

WY and YS performed the data analysis and drafted the manuscript. YZ, HG, and XL designed the study. WY, JF, ZW, LL, XC, JL, YT, DD, DW, and SG contributed to data collection. XL, HG, HJ, and PL performed the revision of the current literature. All authors contributed to the article and approved the submitted version.

FUNDING

This study has received funding by the National Key Research and Development Program of China (2017YFB1304400), the National Natural Science Foundation of China (81901197 and 82171289).

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol.* (2011) 10:626–36. doi: 10.1016/S1474-4422(11)70109-0
2. Weir B. Unruptured intracranial aneurysms: a review. *J Neurosurg.* (2002) 96:3–42. doi: 10.3171/jns.2002.96.1.0003
3. Brown RD Jr, Broderick JP. Unruptured intracranial aneurysms: epidemiology, natural history, management options, and familial screening. *Lancet Neurol.* (2014) 13:393–404. doi: 10.1016/S1474-4422(14)70015-8
4. Li MH, Chen SW, Li YD, Chen YC, Cheng YS, Hu DJ, et al. Prevalence of unruptured cerebral aneurysms in Chinese adults aged 35–75 years: a cross-sectional study. *Ann Intern Med.* (2013) 159:514–21. doi: 10.7326/0003-4819-159-8-201310150-00004

5. Investigators UJ, Morita A, Kirino T, Hashi K, Aoki N, Fukuhara S, et al. The natural course of unruptured cerebral aneurysms in a Japanese cohort. *N Engl J Med*. (2012) 366:2474–82. doi: 10.1056/NEJMoa1113260
6. Darsaut TE, Findlay JM, Magro E, Kotowski M, Roy D, Weill A, et al. Surgical clipping or endovascular coiling for unruptured intracranial aneurysms: a pragmatic randomized trial. *J Neurol Neurosurg Psychiatry*. (2017) 88:663–8. doi: 10.1136/jnnp-2016-315433
7. Dammert S, Krings T, Moller-Hartmann W, Ueffing E, Hans FJ, Willmes K, et al. Detection of intracranial aneurysms with multislice CT: comparison with conventional angiography. *Neuroradiology*. (2004) 46:427–34. doi: 10.1007/s00234-003-1155-1
8. McKinney AM, Palmer CS, Truwit CL, Karagulle A, Teksam M. Detection of aneurysms by 64-section multidetector CT angiography in patients acutely suspected of having an intracranial aneurysm and comparison with digital subtraction and 3D rotational angiography. *AJNR Am J Neuroradiol*. (2008) 29:594–602. doi: 10.3174/ajnr.A0848
9. Sailer AM, Wagemans BA, Nelemans PJ, Graaf Rde, van Zwam WH. Diagnosing intracranial aneurysms with MR angiography: systematic review and meta-analysis. *Stroke*. (2014) 45:119–26. doi: 10.1161/STROKEAHA.113.003133
10. Saitoh H, Hayakawa K, Nishimura K, Okuno Y, Teraura T, Yumitori K, et al. Rerupture of cerebral aneurysms during angiography. *AJNR Am J Neuroradiol*. (1995) 16:539–42.
11. Wang H, Li W, He H, Luo L, Chen C, Guo Y. 320-detector row CT angiography for detection and evaluation of intracranial aneurysms: comparison with conventional digital subtraction angiography. *Clin Radiol*. (2013) 68:e15–20. doi: 10.1016/j.crad.2012.09.001
12. Thompson BG, Brown RD Jr, Amin-Hanjani S, Broderick JP, Cockroft KM, Connolly ES Jr, et al. Guidelines for the management of patients with unruptured intracranial aneurysms: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. (2015) 46:2368–400. doi: 10.1161/STR.0000000000000070
13. Tominari S, Morita A, Ishibashi T, Yamazaki T, Takao H, Murayama Y, et al. Prediction model for 3-year rupture risk of unruptured cerebral aneurysms in Japanese patients. *Ann Neurol*. (2015) 77:1050–9. doi: 10.1002/ana.24400
14. Greving JP, Wermer MJ, Brown RD Jr, Morita A, Juvela S, Yonekura M, et al. Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies. *Lancet Neurol*. (2014) 13:59–66. doi: 10.1016/S1474-4422(13)70263-1
15. Forget TR Jr, Benitez R, Veznedaroglu E, Sharan A, Mitchell W, Silva M, et al. A review of size and location of ruptured intracranial aneurysms. *Neurosurgery*. (2001) 49:1322–5; discussion 1325-6. doi: 10.1097/00006123-200112000-00006
16. Molyneux AJ, Kerr RS, Yu LM, Clarke M, Sneade M, Yarnold JA, et al. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping vs. endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomized comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion. *Lancet*. (2005) 366:809–17. doi: 10.1016/S0140-6736(05)67214-5
17. Molyneux A, Kerr R, Stratton I, Sandercock P, Clarke M, Shrimpton J, et al. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping vs. endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomized trial. *Lancet*. (2002) 360:1267–74. doi: 10.1016/S0140-6736(02)11314-6
18. Wiebers DO, Whisnant JP, Huston J III, Meissner I, Brown RD Jr, Piepgras DG, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet*. (2003) 362:103–10. doi: 10.1016/S0140-6736(03)13860-3
19. Korja M, Lehto H, Juvela S. Lifelong rupture risk of intracranial aneurysms depends on risk factors: a prospective Finnish cohort study. *Stroke*. (2014) 45:1958–63. doi: 10.1161/STROKEAHA.114.005318
20. Lee EJ, Lee HJ, Hyun MK, Choi JE, Kim JH, Lee NR, et al. Rupture rate for patients with untreated unruptured intracranial aneurysms in South Korea during 2006–2009. *J Neurosurg*. (2012) 117:53–9. doi: 10.3171/2012.3.JNS111221
21. Malhotra A, Wu X, Forman HP, Grossetta Nardini HK, Matouk CC, Gandhi D, et al. Growth and rupture risk of small unruptured intracranial aneurysms: a systematic review. *Ann Intern Med*. (2017) 167:26–33. doi: 10.7326/M17-0246
22. Liu Q, Jiang P, Jiang Y, Ge H, Li S, Jin H, et al. Prediction of aneurysm stability using a machine learning model based on pyradiomics-derived morphological features. *Stroke*. (2019) 50:2314–21. doi: 10.1161/STROKEAHA.119.025777
23. Kang H, Ji W, Qian Z, Li Y, Jiang C, Wu Z, et al. Aneurysm characteristics associated with the rupture risk of intracranial aneurysms: a self-controlled study. *PLoS ONE*. (2015) 10:e0142330. doi: 10.1371/journal.pone.0142330
24. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, et al. A study on the application and use of artificial intelligence to support drug development. *Clin Ther*. (2019) 41:1414–26. doi: 10.1016/j.clinthera.2019.05.018
25. Azuaje F. Artificial intelligence for precision oncology: beyond patient stratification. *NPJ Precis Oncol*. (2019) 3:6. doi: 10.1038/s41698-019-0078-1
26. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. (2017) 2:230–43. doi: 10.1136/svn-2017-000101
27. Boland GW, Guimaraes AS, Mueller PR. The radiologist's conundrum: benefits and costs of increasing CT capacity and utilization. *Eur Radiol*. (2009) 19:9–11; discussion 12. doi: 10.1007/s00330-008-1159-7

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.

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fneur.2022.912984

OPEN ACCESS

EDITED BY
Baofeng Gao,
Beijing Institute of Technology, China

REVIEWED BY
Luis Rafael Moscote-Salazar,
Latinamerican Council of Neurocritical Care (CLaNi), Colombia
Zhongbin Tian,
Southern Medical University, China
Han Cong,
Fifth Medical Center of the PLA General Hospital, China

*CORRESPONDENCE
Peng Liu
skeletonliu@sina.com

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

SPECIALTY SECTION
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

RECEIVED 05 April 2022
ACCEPTED 15 August 2022
PUBLISHED 06 September 2022

CITATION
Wei D, Deng D, Gui S, You W, Feng J, Meng X, Chen X, Lv J, Tang Y, Chen T and Liu P (2022) Machine learning to predict in-stent stenosis after Pipeline embolization device placement. *Front. Neurol.* 13:912984. doi: 10.3389/fneur.2022.912984

Machine learning to predict in-stent stenosis after Pipeline embolization device placement

Dachao Wei[1†], Dingwei Deng[1†], Siming Gui[1†], Wei You[1], Junqiang Feng[1], Xiangyu Meng[1], Xiheng Chen[1], Jian Lv[1], Yudi Tang[1], Ting Chen[2] and Peng Liu[1,3*]

[1]Department of Interventional Neuroradiology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China, [2]School of Biomedical Engineering, Capital Medical University, Beijing, China, [3]Department of Interventional Neuroradiology, Beijing Tiantan Hospital, Beijing, China

Background: The Pipeline embolization device (PED) is a flow diverter used to treat intracranial aneurysms. In-stent stenosis (ISS) is a common complication of PED placement that can affect long-term outcome. This study aimed to establish a feasible, effective, and reliable model to predict ISS using machine learning methodology.

Methods: We retrospectively examined clinical, laboratory, and imaging data obtained from 435 patients with intracranial aneurysms who underwent PED placement in our center. Aneurysm morphological measurements were manually measured on pre- and posttreatment imaging studies by three experienced neurointerventionalists. ISS was defined as stenosis rate >50% within the PED. We compared the performance of five machine learning algorithms (elastic net (ENT), support vector machine, Xgboost, Gaussian Naïve Bayes, and random forest) in predicting ISS. Shapley additive explanation was applied to provide an explanation for the predictions.

Results: A total of 69 ISS cases (15.2%) were identified. Six predictors of ISS (age, obesity, balloon angioplasty, internal carotid artery location, neck ratio, and coefficient of variation of red cell volume distribution width) were identified. The ENT model had the best predictive performance with a mean area under the receiver operating characteristic curve of 0.709 (95% confidence interval [CI], 0.697–0.721), mean sensitivity of 77.9% (95% CI, 75.1–80.6%), and mean specificity of 63.4% (95% CI, 60.8–65.9%) in Monte Carlo cross-validation. Shapley additive explanation analysis showed that internal carotid artery location was the most important predictor of ISS.

Conclusion: Our machine learning model can predict ISS after PED placement for treatment of intracranial aneurysms and has the potential to improve patient outcomes.

KEYWORDS

machine learning, flow diverter, Pipeline embolization device, complication, endovascular treatment, intracranial aneurysm

Introduction

Flow diversion are widely used in the treatment of intracranial aneurysms. Among the various available devices, the Pipeline embolization device (PED; Medtronic, Dublin, Ireland) is the most widely studied. The PED was initially developed and approved for treatment of large and giant aneurysms located on the internal carotid artery (ICA) from the petrous to the superior hypophyseal segments (1). Owing to its high occlusion rate and satisfactory safety profile, PED use has been expanded to treat ruptured and unruptured saccular and non-saccular aneurysms of the anterior communicating, middle cerebral, vertebrobasilar, and posterior inferior cerebellar arteries (2–8). In-stent stenosis (ISS) is a common complication of PED placement and has been defined as intimal hyperplasia within the stent that appears as an unfilled contrast space between the contrast filled vascular cavity and stent on digital subtraction angiography (7). However, long-term complications of PED placement are not well understood. Previous studies have reported that most patients with ISS are asymptomatic and that ISS usually gradually improves; however, it may worsen (8–14). In addition, ISS may result in hemiplegia (14, 15) or even death (11) and cause decreased blood flow velocity (16). Considering that severe stenosis can progress to vascular occlusion and weaken the compensatory ability of the cerebral vasculature, its potential harm cannot be ignored. Therefore, the pathogenesis and predictors of ISS should be studied to improve long-term outcomes.

Previous retrospective studies have shown that balloon angioplasty (17), current smoking (18), prior cerebrovascular stenosis (18), dual antiplatelet therapy non-compliance (13), and anterior circulation location (13) are risk factors for ISS. Increasing year of treatment within the study period was also a risk factor in one study (17). Protective factors include increasing age (17), previous endovascular treatment (17), and statin use (11). However, a comprehensive ISS prediction model has not been developed for patients undergoing PED placement. This study aimed to establish a feasible, effective, and reliable ISS prediction model based on patient clinical and imaging characteristics using machine learning methods. Application of such a model can identify patients at high risk for ISS and enable close follow-up, which should improve long-term outcomes.

Materials and methods

Study population

Data for patients treated with flow diverters in the Department of Interventional Neuroradiology, Beijing Tiantan Hospital between January 2015 and October 2020 were retrospectively collected. Only patients treated using the PED who had at least one angiographic follow-up were eligible for study inclusion. In our center, patients scheduled for implantation of PED was administered with aspirin (100 mg) and clopidogrel (75 mg) for at least 5 days prior to the procedure. And the duration of dual antiplatelet therapy ranged from 3 to >6 months after procedure, and a combination of aspirin (100 mg/day) and clopidogrel (75 mg/day) was the most common antiplatelet regimen. We excluded patients who had experienced subarachnoid hemorrhage within 1 month prior to PED placement and those whose imaging studies before or after treatment were not available. Institutional review board approval was obtained. The requirement for informed consent was waived because the study was retrospective in nature and all data were deidentified. A study flow chart of patient selection is illustrated in Figure 1.

Data collection

We collected and recorded clinical and laboratory data from the electronic medical records and reviewed imaging studies (digital subtraction angiography, computed tomography angiography, magnetic resonance angiography) performed before and after treatment. Perioperative laboratory data included data from 14 days before to 14 days after treatment. Imaging follow up was performed 6 and 12 months after treatment and every year thereafter for 5 years. If a laboratory parameter had multiple recordings, the mean value was recorded. Aneurysm morphological parameters, including maximum diameter, neck diameter, maximum height, perpendicular height, aneurysm width, aspect ratio, size ratio, height/width ratio, neck ratio, and bottleneck factor, were manually measured by three experienced neurointerventionalists according to previously published studies (19, 20) (Supplementary Table 1). Parent artery diameter, proximal parent artery diameter, and distal parent artery diameter (defined as the minimum diameter of the parent artery at the aneurysm neck, 1.5 × parent artery diameter upstream from the neck, and 1.5 × parent artery diameter downstream from the neck, respectively) were measured manually at the same point in the imaging studies before and after treatment (21). Stenosis rate was calculated according to the formula:

$$Stenosis\ Rate\ (SR) = 1 - \frac{parent\ artery\ diameter\ at\ certain\ follow\ up\ (D_x)}{intraopeartive\ parent\ artery\ diameter\ (D_0)}$$

ISS was defined as stenosis rate >50% within the PED. ISS was graded as mild (50–74%), severe (75–99%), or occlusion (100%). Aneurysm occlusion was graded according to the O'Kelly-Marotta (OKM) grading scale (22), which is based on the degree of aneurysmal filling: Total filling, subtotal filling, entry remnant, or no filling.

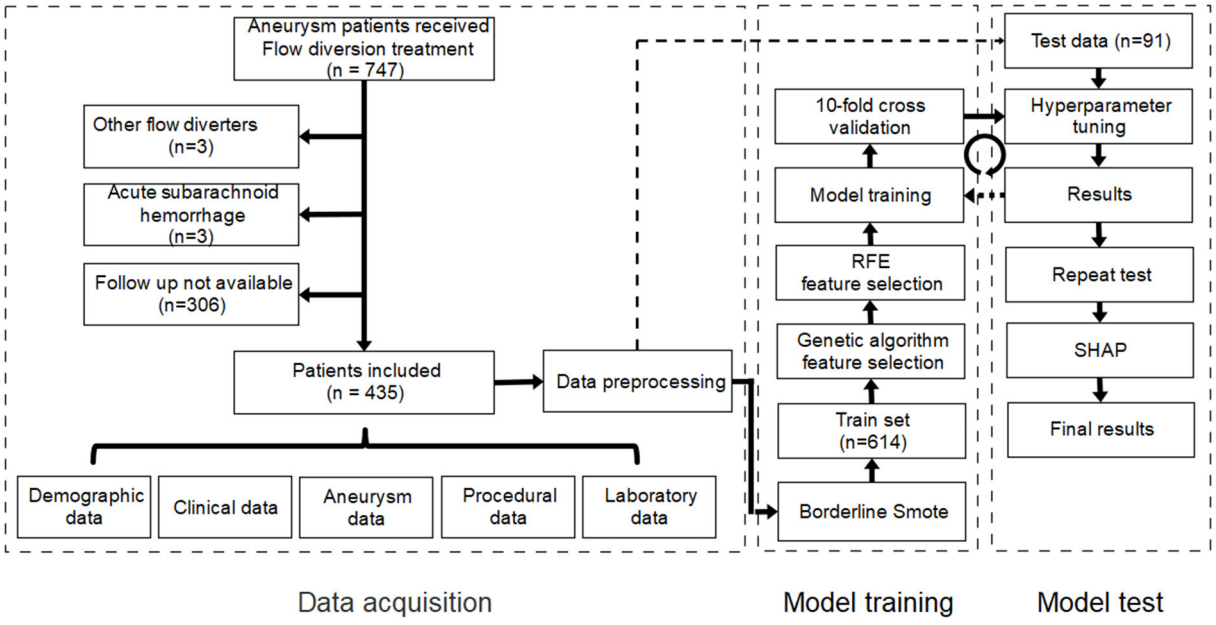


FIGURE 1
Study flow chart of patient selection and establishment of the machine learning model. RFE, recursive feature elimination; SHAP, Shapley additive explanation.

Data preprocessing

Among the 122 variables recorded, 92 were included for analysis after excluding those in which >30% of values were missing (Supplementary Table 2). Missing values were imputed using the random forest method in the missingpy package (version 0.2.0). Continuous variables were standardized using z-score transformation. Categorical variables were binarized. Multicategorical variables were converted into binary variables using one-hot encoding.

The processed dataset was randomly stratified into training (80%) and test (20%) sets. A bias toward negative cases was present because of the scarcity of patients with ISS. Therefore, borderline-SMOTE was applied to the training set using the imblearn package (version 0.8.0). This technique can generate synthetic data from the minority class (patients with ISS) to achieve balance of negative and positive cases (23). After application of borderline-SMOTE, the training set was expanded to 614 cases (307 stenosis cases).

Feature selection and model training

We applied and compared five popular machine learning models: elastic net (ENT), support vector machine (SVM), Xgboost (XGB), Gaussian Naïve Bayes (GNB), and random forest (RF) with traditional logistics regression (LR) using the open-source machine learning library scikit-learn (version 0.24.1). Before model training, genetic algorithm (GA) and recursive feature elimination (RFE) were each applied to the training set to identify the best combination of features. Then, 10-fold cross validation and grid search were used in model training to determine the optimal hyperparameters of each model. The performance of the machine learning models was evaluated using sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) in the test set. The flow chart for model training and testing is shown in Figure 1.

After model training and testing, we applied Monte Carlo cross-validation (MCCV) to verify the efficacy of the machine learning model again. The dataset was randomly divided into test and training sets and the training and testing were repeated 100 times. Sensitivity, specificity, AUC-ROC, maximum Youden index, and threshold at maximum Youden index in each loop were recorded. Mean sensitivity, mean specificity, and mean AUC-ROC were calculated to determine model performance. Mean value of maximum Youden index in each loop was calculated and determined as the optimal threshold.

Model explanation

The Shapley additive explanation (SHAP) algorithm (version 0.39.0) was used to address interpretability problems associated with machine learning models. Based on game theory, SHAP connects optimal credit allocation with local explanations

using the classic Shapley values. SHAP can simultaneously provide local and global model interpretation (24).

Statistical methods

Statistical analyses were performed using Python (version 3.8.8). Categorical variables are expressed as numbers with percentage. Continuous variables with normal distribution are expressed as means ± standard deviation; those with skewed distribution are expressed as medians with interquartile range (IQR). Normality was tested using the Shapiro–Wilk test. One-way analysis of variance was used to compare Monte Carlo cross-validation between the machine learning models. The *post hoc* Tukey honestly significant difference (HSD) test was applied to identify where the differences lay. The highest Youden's index was used to define the optimal cut-off value. The mean value of the optimal cut-off value was used to differentiate low and high stenosis risk. The association between stenosis risk and time after procedure was assessed using Cox regression. The log-rank test was then used to compare Kaplan–Meier curves. Two-tailed $P \leq 0.05$ was considered significant.

Results

Study population and stratified random sampling

Based on our inclusion criteria, 435 patients were finally enrolled. Two hundred and eighty-nine (66.4%) were female. Median age was 54 years (IQR, 47–61). Average body mass index (BMI) was 24.9 (IQR, 22.7–26.7). Sixty-seven patients had BMI >28. One hundred and eighty-four patients (42.3%) had a history of hypertension; 19 (4.4%) had a history of subarachnoid hemorrhage. Seventy-one patients were current or former smokers. Ninety-three aneurysms were non-saccular. Aneurysm location was ICA in 335, vertebral artery in 86, basilar artery 12, middle cerebral artery in 10, and other in 10. Average aneurysm size and neck width were 12.97 ± 8.17 mm and 8.98 ± 6.24 mm, respectively. As of July 2021, 69 ISS cases (15.2%) had been identified; follow-up was available in 66. Among these, 20 (30.3%) were symptomatic. Symptoms included moderate to severe headache (9/20), dizziness or vertigo (5/20), contralateral limb movement disorder (3/20), visual impairment (2/20), neurological deficit (1/20), visual field defect (1/20), and cognitive impairment (1/20). Poor outcome (modified Rankin scale score ≥3) was experienced by five patients (7.6%): one ocular motility disorder, two ISS-related deaths, and two deaths unrelated to ISS (one aneurysm rupture and one acute myocardial infarction).

Random stratification of the cohort resulted in placement of 614 patients (307 stenosis cases) in the training set and 91 (14 stenosis cases) in the test set.

Feature selection

To find the best combination of characteristics, a GA-based program was developed and used; three iterations were performed over the 92 variables in the training set to yield nine predictors (age, obesity, balloon angioplasty, operation duration, size ratio, neck ratio, ICA location, platelet-large cell ratio, and red cell volume distribution width [RDW-CV]). Then we applied the RFE algorithm to the training set and identified 12 predictors (age, height, weight, BMI, obesity, recurrent aneurysm, balloon angioplasty, aneurysm morphology, bifurcation location, ICA location, neck ratio, RDW-CV). Finally, we used the six common features (age, obesity, balloon angioplasty, ICA location, neck ratio, and RDW-CV) in GA and RFE to train the model.

Cross validation and hyperparameter tuning

After 10-fold cross validation and hyperparameter tuning, the best hyperparameters were identified. Model performance is illustrated in Figures 2A,B. In the training set, the XGB model had the highest mean AUC-ROC (0.899; 95% confidence interval [CI], 0.897–0.902), followed by the RF model (0.870; 95% CI, 0.868–0.871), SVM model (0.778; 95% CI, 0.775–0.780), ENT model (0.773; 95% CI, 0.769–0.776), and GNB model (0.772; 95% CI, 0.768–0.775). In the validation set, the XGB model also had the best mean AUC-ROC (0.881; 95% CI, 0.861–0.900), followed by the RF model (0.852; 95% CI, 0.831–0.874), SVM model (0.769; 0.742–0.797), ENT model (0.761; 95% CI, 0.733–0.790), and GNB model (0.761; 95% CI, 0.736–0.785). Then, we tested the models in the test set (Figure 2C). We also tested the performance of logistics regression (LR) (Figure 2D). Though the performance was inferior in cross validation, the ENT model had the highest AUC-ROC in the test set (0.740), followed by the RF model (0.709), SVM model (0.664), XGB model (0.630) and GNB model (0.582). LR had an AUC-ROC of 0.697, which was lower than ENT and RF. The confusion matrix was shown in Table 1. ENT model is a combination of lasso regression and ridge regression, which add regular terms to logistics regression to avoid overfitting. Given that, we believed that ENT model is better than LR and can represent the performance of LR.

To exclude the influence of randomness in the process of assigning patients to the training and test sets, we applied Monte Carlo cross-validation and recorded the AUC-ROC, best Youden index, thresholds at best Youden index, and

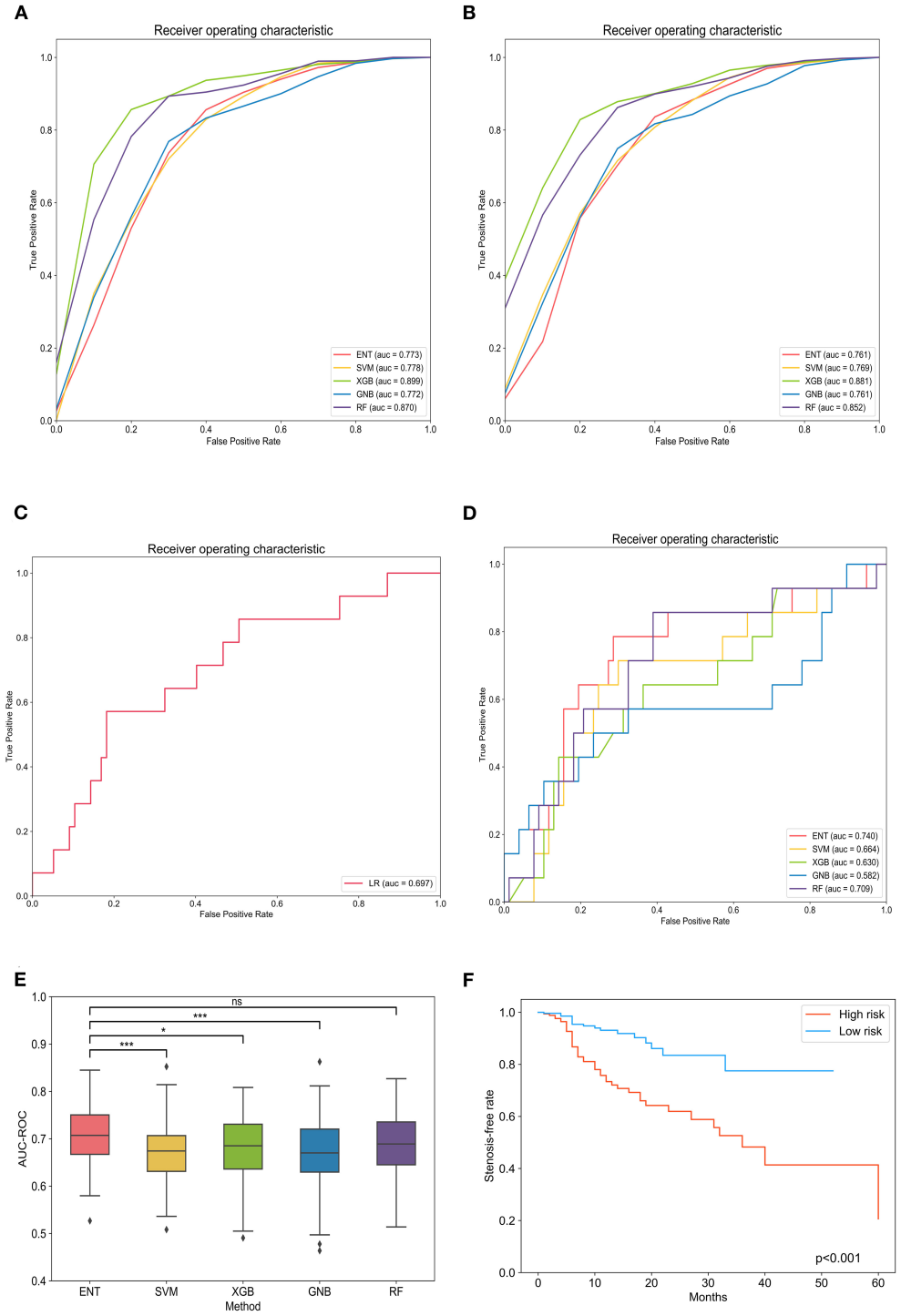


FIGURE 2
Evaluation of machine learning model performance in the training, validation, and test sets. **(A)** Comparison of the area under the receiver operating curve of different models in the training set. **(B)** Comparison of the area under the receiver operating curve of different models in the validation set. **(C)** Comparison of the area under the receiver operating curve of different models in the test set. **(D)** The receiver operating curve of logistics regression. **(E)** Box plot of model area under the receiver operating curves in each loop. *Tukey honestly significant difference (HSD) test $p < 0.05$ between the models; ***Tukey HSD test $p < 0.005$ between the models; ^{ns}Tukey HSD test $p > 0.05$ between the models. **(F)** Kaplan–Meier curves of in-stent stenosis rates for high-risk patients (predicted value > optimal threshold) and low-risk patients (predicted value < optimal threshold). ENT, elastic net; SVM, support vector machine; XGB, Xgboost; GNB, Gaussian Naïve Bayes; RF, random forest; LR, logistics regression.

corresponding sensitivity and specificity in each loop. The ENT model remained the optimal model (0.709; 95% CI, 0.697–0.721) with a mean sensitivity of 77.9% (95% CI, 75.1%−80.6%) and specificity of 63.4% (95% CI, 60.8%−65.9%), followed by the RF model (0.687; 95% CI, 0.674–0.700), XGB model (0.680; 95% CI, 0.668–0.693), GNB model (0.675; 95% CI, 0.661–0.689), and SVM model (0.670; 95% CI, 0.657–0.683; Table 2). One-way analysis of variance and Tukey HSD multiple comparison showed that the ENT model's mean AUC-ROC significantly outperformed the SVM model, XGB model, and GNB model ($p = 0.001$, $p = 0.018$, $p = 0.003$, respectively); however, the mean AUC-ROC did not significantly differ between the ENT and RF models (p = 0.131; Figure 2E).

Model explanation

Spearman correlation testing showed a significant positive correlation between the predicted scores and degree of stenosis ($r = 0.418$, $p < 0.001$). The ENT model was applied to all patients to obtain predicted scores. All patients were grouped according to risk (low-risk and high-risk groups) according to the optimal threshold determined in Monte Carlo cross-validation of Cox regression analysis. Cox regression showed that ISS risk was significantly higher in the high-risk group than the low-risk group (hazard ratio 3.41; 95% CI, 2.03–5.73, $p < 0.001$; Figure 2F).

Next, we used SHAP analysis to interpret the ENT model. Figure 3A shows the importance of the different variables. ICA location had the greatest influence on the model, followed by balloon angioplasty, neck ratio, obesity, RDW-CV, and age. Figure 3B shows the influence of feature values on model prediction. The X-axis represents the influence on the model (SHAP value), the right of the X-axis represents the positive influence, and the left of the X-axis represents the negative influence. The color of the point represents the value of the feature: red represents high feature value and blue represents low feature value. Therefore, balloon angioplasty and increasing neck ratio are risk factors for ISS, while ICA location, obesity, increasing RDW-CV, and increasing age are protective factors. Figure 3C shows the interpretation of SHAP analysis for two individual patients. Case 1 is a patient without ISS in whom the model correctly predicted no stenosis. The influence of various factors on model prediction is shown in the figure. Case 2 is a patient with ISS in whom the model correctly predicted stenosis.

Discussion

We developed a machine learning-based prediction model that can predict ISS in intracranial aneurysm patients who undergo PED placement. Six factors predict ISS: ICA

TABLE 1 Confusion matrix of models in the test set.

	Model name	ENT		SVM		XGB		GNB		RF	
		Stenosis	No stenosis	Stenosis	No stenosis	Stenosis	No stenosis	Stenosis	No stenosis	Stenosis	No stenosis
True value	Stenosis	9	5	6	8	7	7	9	5	6	8
	No stenosis	20	57	20	57	20	57	36	41	14	63

ENT, elastic net; SVM, support vector machine; XGB, Xgboost; GNB, Gaussian Naïve Bayes; RF, random forest. Green: the prediction of models is consistent with the true value. Red: the prediction of models is inconsistent with the true value.

10.3389/fneur.2022.912984

TABLE 2 Comparison of model performance in the training, validation, and test sets.

Model name	Training			Validation			Test		Monte Carlo cross-validation									
	Mean AUC	95% CI		Mean AUC	95% CI		AUC	Accuracy	Mean AUC	95% CI		Mean sensitivity	95% CI		Mean specificity	95% CI		Mean accuracy	95% CI	
		Low	High		Low	High				Low	High		Low	High		Low	High		Low	High
ENT	0.773	0.769	0.776	0.761	0.733	0.790	0.740	0.725	0.709	0.697	0.721	0.779	0.751	0.806	0.634	0.608	0.659	0.666	0.646	0.686
SVM	0.778	0.775	0.780	0.769	0.742	0.797	0.664	0.692	0.670	0.657	0.683	0.758	0.733	0.782	0.624	0.599	0.649	0.641	0.623	0.658
XGB	0.899	0.897	0.902	0.881	0.861	0.900	0.630	0.703	0.680	0.668	0.693	0.719	0.691	0.747	0.662	0.634	0.690	0.669	0.649	0.689
GNB	0.772	0.768	0.775	0.761	0.736	0.785	0.582	0.549	0.675	0.661	0.689	0.673	0.644	0.702	0.707	0.683	0.732	0.699	0.683	0.716
RF	0.870	0.868	0.871	0.852	0.831	0.874	0.709	0.769	0.687	0.674	0.700	0.722	0.693	0.751	0.666	0.640	0.692	0.651	0.633	0.669

AUC, area under the curve; CI, confidence interval; ENT, elastic net; SVM, support vector machine; XGB, Xgboost; GNB, Gaussian Naïve Bayes; RF, random forest.

location, balloon angioplasty, neck ratio, obesity, RDW-CV, and age. Among the five machine learning models, the ENT model had best performance as measured by AUC-ROC, sensitivity, and specificity. Moreover, the result of Monte Carlo cross-validation strongly demonstrated the efficacy and robustness of the machine learning model. We also found a positive correlation between predicted scores and ISS grade. Using the optimum threshold from Monte Carlo cross-validation, we stratified patients according to risk of ISS and showed that the model's risk stratification was accurate. Finally, we utilized SHAP analysis to perform explanations for the machine learning model.

To our knowledge, this is the first prediction model to predict ISS in patients with intracranial aneurysms treated using a flow diverter. ISS is a common complication of flow diverter placement. Sweid et al. reported a 6.3% incidence and noted that ISS was the most common complication (17). A meta-analysis reported an 8.8% incidence (8). In our study, ISS caused symptoms in 30.3% of affected patients and 7.6% experienced a poor outcome. However, most had a good outcome and most patients with ISS were asymptomatic. In addition, ISS in most patients remain stable or even improved. These findings are consistent with previous studies (8–14). Reversible stenosis may be associated with thrombosis (25). Flores-Milan et al. reported an ISS-related death from a stroke secondary to cerebral artery occlusion (11). It remains unclear whether delayed thrombosis, ISS, and patient symptoms are related.

In view of the high incidence and potential harms of ISS, predicting its occurrence, identifying risk factors, and stratifying patients according to risk are necessary to enable better patient care and prevent complications. The ability to predict ISS would enable preoperative evaluation of postoperative risk, which would assist treatment decision making. Furthermore, in patients with low risk of ISS, unnecessary follow-up could be avoided, while high-risk patients would be closely observed and treated appropriately to reduce the risk of acute ischemic complications.

In contrast with the traditional and regular machine-learning based prediction models, our model has several advantages, namely identification of six ISS predictors, use of Borderline-SMOTE in model training, and use of feature selection. We identified six predictors of ISS: ICA location, obesity, increasing RDW-CV, and increasing age were protective factors, while balloon angioplasty and increasing neck ratio were risk factors. Predictors found in previous studies are consistent with ours. Brinjikji et al. (26) found a trend toward higher ISS rates among younger patients; all ISS cases in their study occurred in patients under 50 years of age (2/793). Sweid et al. (17) reported that increasing age is negatively associated with ISS (odds ratio 0.9; $p = 0.02$). Higher rates of ISS in younger individuals have also been reported in stent-assisted coiling and coronary artery stenting studies (27–29);

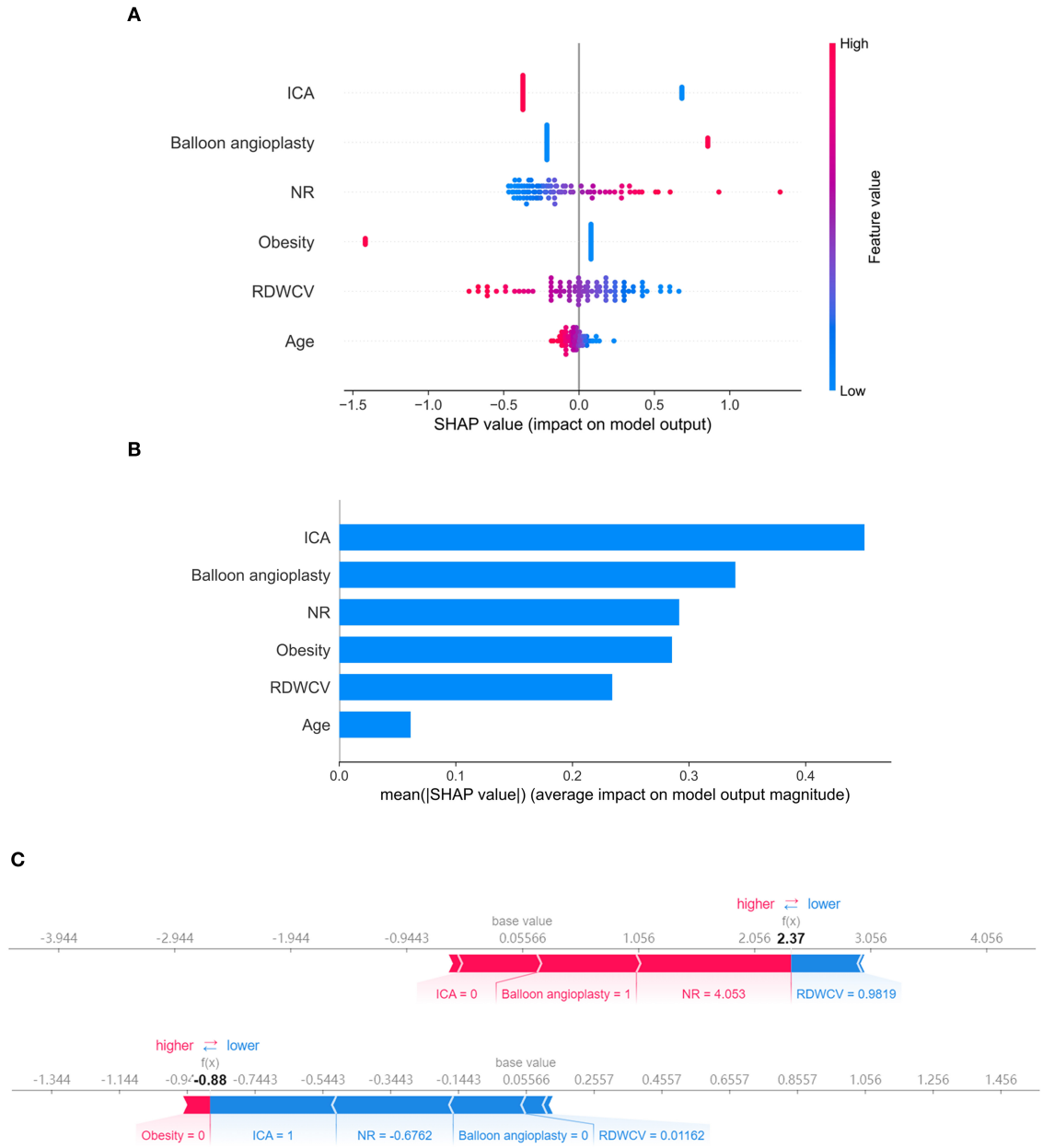


FIGURE 3
Shapley additive explanation (SHAP) analysis of the elastic net (ENT) model. **(A)** Association between the SHAP value and feature value. **(B)** Feature importance (mean |SHAP value|) of each predictor. **(C)** Two ENT model prediction examples. ICA, internal carotid artery; NR, neck ratio; RDWCV, coefficient of variation of red cell volume distribution width.

these higher rates have been attributed to more intense intimal hyperplasia within the device in younger individuals. Sweid et al. (17) also reported balloon angioplasty as an ISS predictor (odds ratio 4.2; $p = 0.03$). John et al. (9) found a higher rate of balloon angioplasty in ISS patients (40 vs. 2%), but they did not conduct statistical inference because of the small number of cases. Balloon angioplasty may result in endothelial damage that induces intimal hyperplasia. This hyperplasia may then progress and eventually cause ISS. In our study, ICA position was negatively associated with ISS, which contradicts the results of Chalouhi et al. (13). The inconsistency may be due to confounding factors. In our cohort, aneurysms in the posterior circulation were mostly fusiform, and those located on the ICA were saccular. Therefore, confounding of location and morphology may have been present. Notedly, Potts et al. (14) reported that fusiform morphology is an ISS predictor

for aneurysms in the anterior circulation. Srinivasan et al. (10) reported similar findings. The fact that fusiform aneurysms may need a longer construct or placement of multiple overlapping devices may explain this, as either may cause more damage to the vascular endothelium. Moreover, fusiform aneurysms tend to have a larger neck width, which take longer to completely endothelialize. Increasing neck ratio was an ISS risk factor in our study, which is in agreement with the findings of Potts et al. (14) Interestingly, obesity (BMI >28) was protective against ISS. In a previous percutaneous coronary intervention meta-analysis, West et al. (30) reported that lower BMI ($p = 0.04$) was associated with restenosis, which is in accordance with our findings. Although obese patients in our study had larger artery diameter than patients with BMI < 28, the difference was not significant (3.90 vs. 3.69 mm; $p = 0.10$, Mann–Whitney U-test). Future studies should elucidate the reason for this finding and explore the relationship between BMI, arterial diameter, and ISS risk. Our study found increasing RDW-CV was a protective factor, which has not been previously reported. We do not yet know the exact mechanism linking RDW-CV and ISS; however, removing RDW-CV from the model will cause a 0.02–0.05 decrease in AUC-ROC. Further work is required to establish the validity of RDW-CV in ISS prediction.

Datasets in classification of diseases or complications are often imbalanced between the numbers of negative and positive cases. Because models based on such datasets may be inaccurate, balancing methods should be implemented. The application of Borderline-SMOTE in our study significantly improved model performance in predicting positive cases; however, it "forged" some positive cases in the strive to balance, which could be controversial in medicine. Therefore, we only used Borderline-SMOTE in the training set; real test data was used to validate the model in model testing.

Feature selection is an important 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. Classical methods of feature selection, such as filter-based methods, which include univariate regression, variance threshold, and maximal information coefficient, have difficulty solving multicollinearity. Therefore, we developed a GA-based feature selection program. A GA simulates the progress of biological evolution. It starts with some chromosomes and individuals (representing a possible combination of features), evaluates the fitness of individuals (AUC-ROC of the validation set), and selects individuals with better fitness to survive, while others will be mutated or crossed over. This process continues until fitness improvement is below the threshold or the maximum number of iterations is reached. In principle, a GA is a random search algorithm. It is possible that it finds a solution that is optimal locally but not globally that is adequate for predicting. We entered 92 variables, iterated over them, and obtained a combination of nine variables. RFE was further used to validate the genetic algorithm results. RFE is a greedy algorithm in essence. It can also achieve a locally optimal solution by removing the most unimportant features repeatedly until the desired number of features is reached. After RFE, there were 12 remaining features, some of which coincided with the GA algorithm, thus verifying the reliability of the GA algorithm. Finally, we used six common features of RFE and GA results to train the model.

Our study has several limitations. First, the study was retrospective in design and conducted in a single center, which may limit the generalizability of our model. A multicenter prospective study is needed in the future for model validation. Second, our dataset had a relatively low number of ISS patients. Although we used Borderline-SMOTE to address this problem, better model performance could be achieved if more ISS cases were available. Third, stenosis measurement was manual and based on different angiographic imaging modalities; therefore, measurement error may have been introduced. However, the mean values of measurements obtained by three different neurointerventionalists were used. In the future, application of deep learning to aneurysm morphology measurement may reduce such errors. Fourth, because of the large number of missing values, we removed all variables in which >30% of the values were missing and used the random forest method to impute missing values in the remaining variables. Fifth, machine learning models are difficult to interpret, which limits their application in medicine. We used SHAP to further illustrate our results. SHAP analysis can provide an explanation for every prediction, which can help clinicians understand model decision making and facilitate application of machine learning models. Sixth, we did not include the length and the diameter of PED in the model because of data deficiency. Longer stent has larger area of contact between the stent and the blood vessels which may result in more damage to the vascular endothelium. Seventh, Exclusion of patients with subarachnoid hemorrhage may weaken the generalization of the results. A subgroup analysis between ruptured aneurysms and unruptured ones may help solve the problem, but we did not have sufficient data.

Conclusion

Our machine learning model can predict ISS after PED placement for treatment of intracranial aneurysms and has the potential to improve patient outcomes.

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 IRB of Beijing Tiantan Hospital Affiliated to Capital Medical University. Written informed consent from the 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

DW, DD, SG, and WY were involved in conception and design of the study. DW, DD, and SG organized the database and performed data analysis. DW drafted the manuscript. WY, JF, XM, XC, JL, YT, TC, and PL critically revised the manuscript and approved its final version. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Key Research and Development Program of China (grant No. 2017YFB1304400) and the Youth Program of the National Natural Science Foundation of China (grant No. 81901197).

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.2022.912984/full#supplementary-material

References

1. Kan P, Siddiqui A, Veznedaroglu E, Liebman KM, Binning MJ, Dumont TM, et al. Early postmarket results after treatment of intracranial aneurysms with the pipeline embolization device: a us multicenter experience. *Neurosurgery.* (2012) 71:1080–7. doi: 10.1227/NEU.0b013e31827060d9

2. Brzezicki G, Rivet DJ, Reavey-Cantwell J. Pipeline Embolization Device for Treatment of High Cervical and Skull Base Carotid Artery Dissections: Clinical Case Series. *J Neurointerv Surg.* (2016) 8:722–8. doi: 10.1136/neurintsurg-2015-011653

3. Natarajan SK, Lin N, Sonig A, Rai AT, Carpenter JS, Levy EI, et al. The safety of pipeline flow diversion in fusiform vertebrobasilar aneurysms: a consecutive case series with longer-term follow-up from a single US center. *J Neurosurg.* (2016) 125:111–9. doi: 10.3171/2015.6.JNS1565

4. Kuhn AL, Kan P, Massari F, Lozano JD, Hou SY, Howk M, et al. Endovascular reconstruction of unruptured intradural vertebral artery dissecting aneurysms with the pipeline embolization device. *J Neurointerv Surg.* (2016) 8:1048. doi: 10.1136/neurintsurg-2015-012028

5. Lauzier DC, Root BK, Kayan Y, Almandoz JED, Osbun JW, Chatterjee AR, et al. Pipeline embolization of distal posterior inferior cerebellar artery aneurysms. *Interv Neuroradiol.* (2021) 27:821–7. doi: 10.1177/15910199211013195

6. Shields LBE, Shields CB, Ghiassi M, Dashti SR, Yao TL, Zhang YP, et al. Pipeline embolization device for treatment of craniocervical internal carotid artery dissections: report of 3 cases. *World Neurosurg.* (2019) 132:106–12. doi: 10.1016/j.wneu.2019.08.183

7. Becske T, Brinjikji W, Potts MB, Kallmes DF, Shapiro M, Moran CJ, et al. Long-term clinical and angiographic outcomes following pipeline embolization device treatment of complex internal carotid artery aneurysms: five-year results of the pipeline for uncoilable or failed aneurysms trial. *Neurosurgery.* (2017) 80:40–8. doi: 10.1093/neuros/nyw014

8. Ravindran K, Salem MM, Enriquez-Marulanda A, Alturki AY, Moore JM, Thomas AJ, et al. Quantitative assessment of in-stent stenosis after pipeline embolization device treatment of intracranial aneurysms: a single-institution series and systematic review. *World Neurosurg.* (2018) 120:e1031–e40. doi: 10.1016/j.wneu.2018.08.225

9. John S, Bain MD, Hui FK, Hussain MS, Masaryk TJ, Rasmussen PA, et al. Long-term follow-up of in-stent stenosis after pipeline flow diversion treatment of intracranial aneurysms. *Neurosurgery.* (2016) 78:862–7. doi: 10.1227/NEU.0000000000001146

10. Srinivasan VM, Mokin M, Duckworth EAM, Chen S, Puri A, Kan P. Tourniquet parent artery occlusion after flow diversion. *J Neurointerv Surg.* (2018) 10:122–6. doi: 10.1136/neurintsurg-2016-012937

11. Flores-Milan G, Pressman E, Peto I, Ren Z, Guerrero WR, Mokin M. Factors Associated with in-stent stenosis after cerebral aneurysm embolization using a pipeline embolization device. *Interv Neuroradiol.* (2021). doi: 10.1177/15910199211066368. [Epub ahead of print].

12. Lauzier DC, Cler SJ, Chatterjee AR, Osbun JW, Moran CJ, Kansagra AP. The value of long-term angiographic follow-up following pipeline embolization of intracranial aneurysms. *J Neurointerv Surg.* (2021). doi: 10.1136/neurintsurg-2021-SNIS.241

13. Chalouhi N, Polifka A, Daou B, Kung D, Barros G, Tjoumakaris S, et al. In-pipeline stenosis: incidence, predictors, and clinical outcomes. *Neurosurgery.* (2015) 77:875–9. doi: 10.1227/NEU.0000000000000908

14. Potts MB, Shapiro M, Zumofen DW, Raz E, Nossek E, DeSousa KG, et al. Parent vessel occlusion after pipeline embolization of cerebral aneurysms of the anterior circulation. *J Neurosurg.* (2017) 127:1333–41. doi: 10.3171/2016.9.JNS152638

15. Fujii S, Fujita K, Yamaoka H, Miki K, Hirai S, Nemoto S, et al. Refractory in-stent stenosis after flow diverter stenting associated with delayed cobalt allergic reaction. *J Neurointerv Surg.* (2021) 14:e4. doi: 10.1136/neurintsurg-2021-017948

16. McDougall CM, Khan K, Saqqur M, Jack A, Rempel J, Derksen C, et al. Ultrasound for the evaluation of stenosis after flow diversion.

J Neurointerv Surg. (2018) 10:297–300. doi: 10.1136/neurintsurg-2017-013049

17. Sweid A, Starke RM, Herial N, Chalouhi N, Das S, Baldassari MP, et al. Predictors of complications, functional outcome, and morbidity in a large cohort treated with flow diversion. *Neurosurgery.* (2020) 87:730–43. doi: 10.1093/neuros/nyz508

18. Luo B, Kang H, Zhang H, Li T, Liu J, Song D, et al. Pipeline embolization device for intracranial aneurysms in a large Chinese cohort: factors related to aneurysm occlusion. *Ther Adv Neurol Disord.* (2020) 13:1756286420967828. doi: 10.1177/1756286420967828

19. Mocco J, Brown RDJr, Torner JC, Capuano AW, Fargen KM, Raghavan ML, et al. aneurysm morphology and prediction of rupture: an international study of unruptured intracranial aneurysms analysis. *Neurosurgery.* (2018) 82:491–6. doi: 10.1093/neuros/nyx226

20. Juvela S, Korja M. Intracranial aneurysm parameters for predicting a future subarachnoid hemorrhage: a long-term follow-up study. *Neurosurgery.* (2017) 81:432–40. doi: 10.1093/neuros/nyw049

21. Dhar S, Tremmel M, Mocco J, Kim M, Yamamoto J, Siddiqui AH, et al. morphology parameters for intracranial aneurysm rupture risk assessment. *Neurosurgery.* (2008) 63:185–96. doi: 10.1227/01.NEU.0000316847.64140.81

22. O'Kelly C J, Krings T, Fiorella D, Marotta TR, A. novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents. *Interv Neuroradiol.* (2010) 16:133–7. doi: 10.1177/159101991001600204

23. Hui H, Wang WY, Mao BH, editors. Borderline-smote: a new over-sampling method in imbalanced data sets learning. In: *Proceedings of the 2005 International Conference on Advances in Intelligent Computing - Volume Part I* (2005).

24. Lundberg S, Lee SI. A unified approach to interpreting model predictions. *Nips.* (2017) 30. doi: 10.48550/arXiv.1705.07874

25. Monteiro A, Lopes DK, Aghaebrahim A, Hanel R. Optical coherence tomography for elucidation of flow-diversion phenomena: the concept of endothelized mural thrombus behind reversible in-stent stenosis in flow-diverters. *Interv Neuroradiol.* (2021):15910199211003432. doi: 10.1177/15910199211003432

26. Brinjikji W, Kallmes DF, Cloft HJ, Lanzino G. Age-related outcomes following intracranial aneurysm treatment with the pipeline embolization device: a subgroup analysis of the intreped registry. *J Neurosurg.* (2016) 124:1726–30. doi: 10.3171/2015.5.JNS15327

27. Turk AS, Levy EI, Albuquerque FC, Pride GLJr, Woo H, Welch BG, et al. Influence of patient age and stenosis location on Wingspan in-stent restenosis. *AJNR Am J Neuroradiol.* (2008) 29:23–7. doi: 10.3174/ajnr.A0869

28. Gao B, Safain MG, Malek AM. Enterprise stenting for intracranial aneurysm treatment induces dynamic and reversible age-dependent stenosis in cerebral arteries. *J Neurointerv Surg.* (2015) 7:297–302. doi: 10.1136/neurintsurg-2013-011074

29. Du R, Zhang RY, Zhang Q, Shi YH, Hu J, Yang ZK, et al. Assessment of the relation between ivus measurements and clinical outcome in elderly patients after sirolimus-eluting stent implantation for De Novo coronary lesions. *Int J Cardiovasc Imaging.* (2012) 28:1653–62. doi: 10.1007/s10554-011-0007-z

30. West NE, Ruygrok PN, Disco CM, Webster MW, Lindeboom WK, O'Neill WW, et al. Clinical and angiographic predictors of restenosis after stent deployment in diabetic patients. *Circulation.* (2004) 109:867–73. doi: 10.1161/01.CIR.0000116750.63158.94

frontiers | Frontiers in Neurology

TYPE Systematic Review
PUBLISHED 08 September 2022
DOI 10.3389/fneur.2022.945813

OPEN ACCESS

EDITED BY
Yuhua Jiang,
Beijing Tiantan Hospital, Capital Medical University, China

REVIEWED BY
Ji Wenjun,
Yulin No. 2 People's Hospital, China
Xiang Zhou,
Second Affiliated Hospital of Soochow University, China

*CORRESPONDENCE
Minyan Zeng
minyan.zeng@adelaide.edu.au

SPECIALTY SECTION
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

RECEIVED 17 May 2022
ACCEPTED 18 August 2022
PUBLISHED 08 September 2022

CITATION
Zeng M, Oakden-Rayner L, Bird A, Smith L, Wu Z, Scroop R, Kleinig T, Jannes J, Jenkinson M and Palmer LJ (2022) Pre-thrombectomy prognostic prediction of large-vessel ischemic stroke using machine learning: A systematic review and meta-analysis. *Front. Neurol.* 13:945813. doi: 10.3389/fneur.2022.945813

Pre-thrombectomy prognostic prediction of large-vessel ischemic stroke using machine learning: A systematic review and meta-analysis

Minyan Zeng[1,2]*, Lauren Oakden-Rayner[1,2,3], Alix Bird[1,2], Luke Smith[1,2], Zimu Wu[4], Rebecca Scroop[3,5], Timothy Kleinig[5,6], Jim Jannes[5,6], Mark Jenkinson[1,7] and Lyle J. Palmer[1,2]

[1]Australian Institute for Machine Learning, University of Adelaide, Adelaide, SA, Australia, [2]School of Public Health, University of Adelaide, Adelaide, SA, Australia, [3]Department of Radiology, Royal Adelaide Hospital, Adelaide, SA, Australia, [4]School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia, [5]Faculty Health and Medical Science, School of Medicine, University of Adelaide, Adelaide, SA, Australia, [6]Department of Neurology, Royal Adelaide Hospital, Adelaide, SA, Australia, [7]Functional Magnetic Resonance Imaging of the Brain Centre, University of Oxford, Oxford, United Kingdom

Introduction: Machine learning (ML) methods are being increasingly applied to prognostic prediction for stroke patients with large vessel occlusion (LVO) treated with endovascular thrombectomy. This systematic review aims to summarize ML-based pre-thrombectomy prognostic models for LVO stroke and identify key research gaps.

Methods: Literature searches were performed in Embase, PubMed, Web of Science, and Scopus. Meta-analyses of the area under the receiver operating characteristic curves (AUCs) of ML models were conducted to synthesize model performance.

Results: Sixteen studies describing 19 models were eligible. The predicted outcomes include functional outcome at 90 days, successful reperfusion, and hemorrhagic transformation. Functional outcome was analyzed by 10 conventional ML models (pooled AUC=0.81, 95% confidence interval [CI]: 0.77–0.85, AUC range: 0.68–0.93) and four deep learning (DL) models (pooled AUC=0.75, 95% CI: 0.70–0.81, AUC range: 0.71–0.81). Successful reperfusion was analyzed by three conventional ML models (pooled AUC=0.72, 95% CI: 0.56–0.88, AUC range: 0.55–0.88) and one DL model (AUC=0.65, 95% CI: 0.62–0.68).

Conclusions: Conventional ML and DL models have shown variable performance in predicting post-treatment outcomes of LVO without generally demonstrating superiority compared to existing prognostic scores. Most models were developed using small datasets, lacked solid external validation, and at high risk of potential bias. There is considerable scope to improve

study design and model performance. The application of ML and DL methods to improve the prediction of prognosis in LVO stroke, while promising, remains nascent.

Systematic review registration: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021266524, identifier CRD42021266524

KEYWORDS

ischemic stroke, large vessel occlusion, endovascular thrombectomy, prognostic prediction, machine learning, deep learning

Introduction

Ischemic stroke caused by large vessel occlusion (LVO) accounts for 24–46% of ischemic stroke cases (1). Endovascular thrombectomy (EVT) is currently the standard care for ischemic stroke patients with occlusion in the anterior cerebral circulation and salvageable brain tissue within 24 h of symptom onset (2). However, despite advances in stroke treatment, the rate of long-term disability/dependency is up to approximately 50% in LVO patients (3). Further, EVT is resource intensive. Better identification of the risks and benefits of intervention may be valuable to optimize patient outcomes and reduce healthcare and societal costs.

To help improve treatment strategies and clinical decision-making, prior studies have investigated pre-treatment predictors of key clinical outcomes following LVO stroke, including comorbidities, clinical examination, and neuroimaging findings (4). A number of prognostic scores using simple linear combinations of these predictors, such as ASPECTS, HIAT, and MR PREDICTS, have been constructed and validated in LVO cohorts treated with EVT (4). However, they may have low clinical utility due to their modest performance in practice (4). Other barriers of their clinical implementation include complexity of scoring and the subjective nature of data acquisition, which are time-dependent with concomitant high inter-observer variability (5, 6). There is a need for a more robust and clinically useful prognostic tool.

Machine learning (ML) techniques are being increasingly applied to clinical tasks (7). These techniques have the potential to handle a large quantity of data and identify latent patterns and complex relationships (8). Deep learning (DL), a newer type of ML technique, can automatically learn useful features at the pixel or voxel level, which is particularly powerful in processing raw medical images (9). DL has shown substantial promise in clinical prognostic prediction based on raw image data (10, 11), and, therefore, may play a role in predicting stroke outcomes—an area characterized by rich neuroimaging datasets.

This systematic review aimed to evaluate the performance, validity, and clinical applicability of published ML-based pre-thrombectomy prognostic models for LVO stroke and to identify key research gaps.

Methods

This systematic review was registered on PROSPERO (12) (ID: CRD42021266524) and conducted in line with the PRISMA guidelines (13).

Eligibility criteria

Publications were eligible for inclusion if the study applied ML and/or DL algorithms to predict clinical outcomes following EVT treatment of LVO stroke. Specifically, the studies were included if: 1) the prediction models were applied to LVO stroke patients treated with EVT; and 2) the study employed ML-based algorithms, such as random forest analysis, naive Bayes classifiers, support vector machines, regression models, and/or various DL algorithms such as convolutional neural networks. Standard regression models without penalization (such as simple logistic regression, linear regression, and cox regression models) were not considered within the scope of this review.

Studies were excluded if: 1) the prediction models included patients with non-LVO stroke such as intracerebral hemorrhage or lacunar stroke; 2) assessment of the model performance was not performed; or 3) the prediction models involved post-EVT information. Conference abstracts, review articles, letters, comments, editorials, and erratum were excluded due to limited information contained.

Search strategies

Full details of the search strategies are shown in Supplementary Table S1. A variety of keywords were selected for literature search after consultation with an academic librarian. Systematic searches were conducted in four databases—PubMed, Embase, Scopus, and Web of Science, from inception until the 18th February 2022. These databases included related computer science conferences and journal papers, except the International Conference on Medical Imaging with Deep Learning (MIDL), so manual searches in MIDL were conducted

to supplement the searches in online databases. Searches were limited to studies published in English.

Study selection

Two reviewers (MZ and ZW) independently conducted study selection and review. After removing duplicates, conference abstracts, narrative reviews, comments, letters, editorial and erratum, the records were screened based on the titles and abstracts, and subsequently assessed by full-text reading. Discrepancies between the two reviewers were resolved by discussion and consultation with a third reviewer (LJP).

Data extraction

Relevant data from the eligible studies were extracted into a pre-specified form independently by two reviewers (MZ and ZW). The data extracted were: 1) year of publication; 2) sample sizes of the training, testing, and external validation cohorts if applicable; 3) demographic characteristics of the study population (age, gender, and ethnicity/place of recruitment); 4) vessel occlusion sites; 5) clinical outcomes assessed; 6) imaging modality used for model development; 7) specific algorithms used; 8) model performance; and 9) model validation. Information related to model development and model performance was restricted to that pertaining to the "best-performing" model. A third reviewer (LJP) resolved any disagreements regarding the extracted information between the two reviewers.

Data synthesis

The model performance was quantified by area under the receiver operating characteristic curve (AUC), an estimation for the discriminative capacity of a model. The AUCs and 95% confidence intervals (CIs) of relevant models were extracted and synthesized. The standard error of each AUC was calculated using the actual positive endpoint and actual negative endpoint based on formula provided in Bradley et al. (14). To make analyses consistent, 95% CIs were calculated based on the information available in the reports using the statistical formula (15): 95% CI = effect size (AUC) $\pm$ 1.96 $\times$ standard error. "Significant" statistical heterogeneity was defined using the Cochran's *Q*-test ($P \leq 0.10$) and the I^2 statistic (>50%) (16). AUCs were pooled in a random-effects model if there was significant heterogeneity suggested by the *Q*-test or I^2. Otherwise, the AUCs were pooled using a fixed-effects model. For adequate statistical power, we used Egger's test with a funnel plot to detect publication bias only when a meta-analysis included more than 10 AUCs and had no statistically substantial heterogeneity suggested by the I^2 or *Q*-test (17, 18). The meta-analyses were conducted using the MedCalc Statistical Software (version 20.0.3).

Risk of bias and reporting quality

Assessment of risk of bias was conducted using the Prediction Model Risk of Bias Assessment Tool (PROBAST) (19). This tool contains 20 questions covering four domains, including participants, predictors, outcomes, and analysis. Assessment of the adherence to reporting standards was conducted using the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD) protocol (20). This checklist contains 22 items (37 points) covering multiple aspects, including title and abstract, backgrounds and objectives, methods, results, discussion, supplementary and funding. In TRIPOD and PROBAST, items related to the details of predictors were not applicable for studies using DL models. This was because "predictors" in DL models are usually each pixel or voxel of an image, which are less likely to be reported in DL models (21). The modified TRIPOD and PROBAST are shown in Supplementary Tables S2, S3.

Results

Search results

A total of 4,116 records were identified in the initial search. After the review of titles and abstracts and the screening of full texts, 16 studies met the inclusion criteria and were included in the systematic review (Figure 1).

Basic characteristics

The basic characteristics of the eligible studies (22–37) are summarized in Supplementary Table S4. The mean or median ages of the study participants ranged from 64.0 to 86.0 years, and the proportion of male participants ranged from 35.0 to 65.9%. Only one US study (24) specifically described the self-reported ethnicity of the patients (63.0–69.0% European ancestry); the other studies reported the place of patient recruitment [USA: 1 (32); Europe: 10 (22, 23, 26–29, 31, 33–35); Asia: 4 (25, 30, 36, 37)]. The training sample sizes ranged widely, from 109 to 1,401. Regarding the testing sample, two studies used hold-out test sets, respectively containing 208 patients (30) and 100 patients (35). The remaining studies performed cross-validation (23–26, 28, 29, 31–34, 36, 37) or bootstrap approach (22, 27). The five studies (23, 29, 31, 34, 35) used data obtained from MR CLEAN Registry (38). Fifteen studies reported the occlusion sites, of which 14 studies (22, 23, 25–36) included patients with anterior

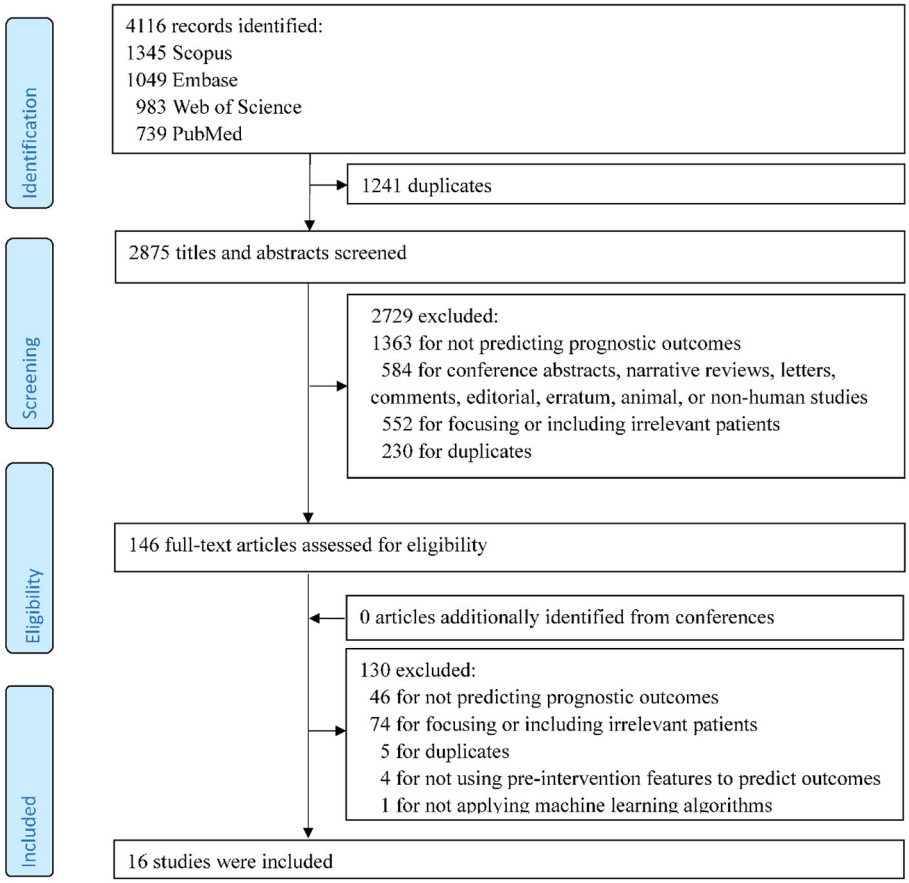


FIGURE 1
Flow chart of study selection.

circulation occlusion and one (24) further included patients with occlusion in the posterior circulation.

Model development

Conventional machine learning algorithms

Details of model development in 12 studies using conventional ML algorithms are shown in Table 1. Tree models (22, 24, 31), random forests (23, 26, 27), and support vector machines (28, 30, 33) were each proposed by three studies, regularized logistic regression by two studies (25, 32), and artificial neural networks by one study (29). To accommodate missing values, two studies used multiple imputation (23, 29) and one used singular imputation (31), while other studies excluded participants with missing data in either predictive or outcome variables (complete-case analysis) (22, 24–28, 30, 32, 33). The number of predictive variables used for model construction varied from 4 (32) to 53 (23). The National Institutes of Health Stroke Scale and age were commonly ranked as the important predictors. All studies conducted internal validation, either by bootstrapping (22, 27), hold-out validation (30), or k-fold cross-validation (23–26, 28, 29, 31–33).

Deep learning algorithms

Table 2 summarizes the model development of DL algorithms in four studies. All studies conducted skull stripping, augmentation, normalization, and imaging resampling (34–37). Two studies (36, 37) additionally labeled regions of interest in the scans. All studies used DL algorithms based on supervised learning (34–37), with one study also using unsupervised learning (auto-encoder) for model pre-training (34). Regarding model architectures, Hilbert et al. (34) used a convolutional auto-encoder to obtain representative imaging features and applied a 2-D ResNet for fine-tuning in successful reperfusion prediction, while the auto-encoder was not used in the best model for functional outcome prediction. The authors utilized structured receptive field kernels (as opposed to learned convolutional kernels) to help prevent overfitting. Samak et al.

TABLE 1 Model development using conventional machine learning algorithms.

References	Model	Outcomes	Missing value	Features	Important feature identified	Validation
Brugnara et al. (22)	Tree model: Gradient boosting decision trees	Good functional outcomes (mRS≤2)	Patients with missing data were excluded	16	Premorbid mRS, baseline acute ischemic volume, NIHSS, onset to imaging time, baseline eASPECTS	Bootstrapping (25 bootstrap sample)
Van et al. (23)	RFA	a. Good functional outcomes (mRS≤2) b. Successful reperfusion (TICI score≥2b)	Patients with missing data of main outcomes were excluded; other variables, multiple imputations	53	Age, NIHSS at baseline, duration of onset to groin puncture, Glasgow Coma Scale, systolic BP at baseline, CRP, creatinine, thrombocyte count, diastolic BP at baseline, baseline ASPECTS, glucose, clot burden score; feature importance for good functional outcomes only: baseline mRS, presence of leukoaraiosis, collateral score; feature importance for successful reperfusion only: occlusion site, hyperdense artery sign, history of AF	Nested cross-validation: 100 repeated random splits; 10-fold cross validation
Alawieh et al. (24)	Tree model (regression tree)	Good functional outcomes (mRS≤2) *	Patients with missing data were excluded	12	Age, gender, race, diabetes, hypertension, hyperlipidemia, arterial fibrillation, preceding intravenous thrombolysis, onset to groin puncture time, NIHSS, baseline mRS, ASPECTS**	10-fold cross-validation
Nishi et al. (25)	RLR	Good functional outcomes (mRS≤2)	Patients with missing data were excluded	16	Care-dependent, age, premorbid mRS, ASPECTS, NIHSS	10-fold cross-validation
Hamann et al. (26)	RFA	Good functional outcomes (mRS≤2)	Patients with missing data were excluded	10	Age, NIHSS at baseline, systolic blood pressure, risk factors (hypertension, diabetes, smoking, previous ischemic event), preceding intravenous thrombolysis, onset to groin puncture time, collateralization status, perfusion value of the medial MCA territory, volume of core, and volume of tissue at risk**	5-fold cross validation
Kerleroux et al. (27)	RFA	Good functional outcomes (mRS≤3)	Patients with missing data were excluded	32	Receiving mechanical thrombectomy, the absence of ICA occlusion, lower HE-I, decreasing age, and the presence of eloquent mismatch within the following regions: the right thalamus, the left thalamus, the left superior longitudinal fasciculus, the left post central gyrus, the left retro-lenticular part of internal capsule, and the left supra marginal gyrus	Bootstrapping
Xie et al. (28)	SVM	Good functional outcomes (mRS≤2)	Patients with missing data were excluded	4	Age, baseline NIHSS score, lesion volume, ischemic percentage in each brain region	Nested cross-validation: 100 repeated random splits; 10-fold cross validation
Ramos et al. (29)	ANN	Poor functional outcomes (mRS≥5)	Multiple imputation	51	Age, collateral, glucose level, NIHSS, and pre-stroke mRS	Nested cross-validation: 10 equally sized splits; 5-fold cross validation
Ryu et al. (30)	SVM	Poor functional outcome (mRS≥4)	Patients with missing data were excluded	6	Age, NIHSS, hypertension, diabetes mellitus, AF, and poor collateral*	Hold-out validation
Kappelhof et al. (31)	Tree model (Decision tree)	Poor functional outcome (mRS≥5)	Singular imputation	6	Age, pre-stroke mRS, start of endovascular thrombectomy, NIHSS at baseline, history of diabetes mellitus, duration of CTA in first hospital to groin puncture*	5-fold cross-validation
Patel et al. (32)	RLR	Successful reperfusion at the first attempt (TICI score≥2b)	Patients with missing data were excluded	4	Clot length, clot perviousness, distance from internal carotid artery, angle between the aspiration catheter and the clot	Nested cross-validation: 100 repeated random splits; 10-fold cross validation
Hofmeister et al. (33)	SVM	Successful reperfusion at the first attempt (TICI score≥2b)	Patients with missing data were excluded	9	Large area low gray level emphasis, gray level variance, large dependence emphasis, short run emphasis, entropy, maximum, run percentage, coarseness, and gray level nonuniformity normalized*	10-fold cross-validation

AF, atrial fibrillation; ASPECTS, The Alberta stroke program early CT score; BP, blood pressure; CRP, C-reactive protein; HE-I, high-eloquence infarct; MCA, middle cerebral artery; mRS, Modified Rankin Scale; ICA, internal carotid artery; NIHSS, The National Institutes of Health Stroke Scale; RFA, random forest analysis; RLR, regularized logistic regression; TICI, thrombolysis in cerebral infarction score; n.a., not available; ANN, artificial neural networks; CTA, computed tomography angiography; mRS, Modified Rankin Scale; SVM, support vector machine; TICI, thrombolysis in cerebral infarction score. *Study u tree model to predict continuous multiclass mRS (0, 1, 2, 3, 4, 5, 6) and also dichotomized multiclass mRS ("good" vs. "poor" function) for model prediction and comparison. **Features used in final model were listed here as feature impo ; analysis was not conducted in the included study.

10.3389/fneur.2022.945813

(35) and Jiang et al. (37) both used a 3-D CNN feature encoder and incorporated imaging and clinical data using metadata fusion technique. The former additionally used self-attention technique (squeeze and excitation modules) in their encoders, while the latter is based on pre-trained Inception V3 encoders. Additionally, the latter built the encoder individually on multiple imaging modalities (Diffusion Weight Imaging [DWI], Mean Transit Time map, and Time To Peak map). Nishi et al. (36) used a U-net for predicting ischemic core lesion segmentation to derive feature representations and used a 2-layer neural network on top of feature representations for fine tuning. Two studies used saliency c-map for imaging feature visualization (34, 36). All four studies excluded patients with missing values in either imaging data or outcome measures. Three studies conducted k-fold cross-validation (34, 36, 37) and one used hold-out validation (35).

Model performance

Conventional machine learning algorithms

Model performance of the 13 conventional ML models was summarized in Table 3. Ten models predicted the functional outcome at 90 days post-stroke defined by the mRS (39) (pooled AUC=0.81, 95% CI: 0.77–0.85, AUC range: 0.68–0.93, Figure 2A). Seven of these models used imaging features selected from computed tomography (CT) (pooled AUC=0.82, 95% CI: 0.78–0.86), and three involved features identified in magnetic resonance imaging (MRI) (pooled AUC=0.77, 95% CI: 0.70–0.85) (Supplementary Figure S1). Three models predicted successful reperfusion defined by the Thrombolysis in Cerebral Infarction Score (pooled AUC=0.72, 95% CI: 0.56–0.88, AUC range: 0.55–0.88; Supplementary Figure S2). Three models were validated (24, 25, 33) in external datasets.

Deep learning algorithms

The six DL models were summarized in Table 4. Good functional outcome defined as mRS≤2 was analyzed in three models (pooled AUC=0.75, 95% CI: 0.70–0.81; Figure 2B), among which two were CT-based (AUC range: 0.71–0.75) and one was MRI-based (AUC: internal, 0.81; external, 0.73). The outcomes predicted in the other three models include: each of the seven mRS points (accuracy=0.35), successful reperfusion (AUC=0.65, 95% CI: 0.62–0.68), and hemorrhage transformation (AUC=0.95, 95% CI: 0.87–1.00). Two models conducted external geographic validation (36, 37).

Risk of bias

Three ML-based studies (23, 29, 31) and one DL-based study (34) were considered at low risk of bias in all domains (Supplementary Table S5). The remaining studies were at high risk of bias in at least one domain (22, 24–28, 30, 32, 33, 35–37). Risk of bias mostly occurred in handling missing data. Risks of bias in other items, including standard outcome definition and internal validation techniques, was also identified.

Reporting quality

All studies were rated as "good" in terms of overall adherence (>70% items reported) (Supplementary Table S6). However, several items remained rarely reported, including sample size calculations, how risk groups were defined, the detailed parameters of the prediction models and how to use the prediction model.

Discussion

The application of ML techniques in prognostic prediction for LVO stroke is evolving. CT images have been more commonly used than MRI images in model development. Most studies used short-term reperfusion and functional outcomes at 90 days post-stroke as the prognostic endpoints. Conventional ML and DL models showed similar performance, but neither significantly outperformed existing prognostic scores. Also, many studies exhibited a high risk of potential bias and few studies adequately reported details of the models developed.

Image data

Most studies selected CT over MRI as the imaging modality, in keeping with clinical practice (40). MRI may offer superior outcome prediction because of more precise measurement of early stroke damage, but its availability, acquisition speed and frequent contraindications have proven formidable barriers to routine use (41). Meanwhile, the performance of CT imaging has been improving over time, reducing the diagnostic precision gap (41). Indeed, our review suggests that MRI did not show superior performance to CT in prognostication, bolstering the rationale for developing CT-based prognostic models.

Predicted outcomes

Our review identified clear gaps regarding the outcomes investigated. The only "long-term" outcome investigated was the mRS score at 90 days. This outcome was analyzed as a binary variable in all studies (dichotomized at two or three for good vs. moderate-to-poor outcome; or at four or five for poor vs. moderate-to-good outcome). However, such dichotomization might be arbitrary and inconsistent, which may have introduced

TABLE 2 Model development using deep learning algorithms.

References	Outcomes	Missing value	Major imaging pre-processing	Model architecture	Feature visualization	Validation
Hilbert et al. (34)	a. Good functional outcome (mRS≤2) b. Successful reperfusion (TICI score≥2b)	Patients with missing data were excluded	a. Brain extraction (50–400 HU) b. Rigid registration to a template c. Computing maximum intensity projection from 3D to 2D scans d. Normalization e. Imaging resampling (368 × 432)	a. Functional outcome: supervised 2D-ResNet architecture with structured receptive field kernels model b. Successful reperfusion: a stacked denoising convolutional auto-encoder (2D-ResNet architecture with structured receptive field kernels) and fine-tuned model	Gradient-weighted Class Activation Mapping	4-fold cross validation
Samak et al. (35)	a. Good functional outcome (mRS≤2) b. Individual mRS scores (0–6)	Patients with missing data were excluded	a. Brain extraction (40–100 HU) b. Data augmentation (flip, rotations, elastic deformations, Gaussian noise) c. Normalization d. Imaging resampling (192x192x32)	a. Multimodal model: image feature encoder, clinical metadata encoder, image and clinical metadata fusion b. 3D-convolutional kernels, attentional block	n.a.	Hold-out validation
Nishi et al. (36)	Good functional outcome (mRS≤2)	Patients with missing data were excluded	a. Brain extraction b. Data augmentation (rotations, translation, spatial scaling) c. Normalization d. ROIs labeling (ischemic core lesion) e. Imaging resampling (128 × 128 × 32)	a. Multi-output model: A U-net segmentation task for imaging feature derivation, a 2-layer neural network for fine-tuning b. 3D-convolutional kernels	Gradient-weighted Class Activation Mapping	5-fold cross validation
Jiang et al. (37)	Hemorrhagic transformation (including HI1, HI2, PH1, and PH2)	Patients with missing data were excluded	a. Brain extraction b. Data augmentation (rotations, spatial scaling) c. ROIs labelling d. Imaging resampling (randomly cropped from ROIs)	a. Multimodal model: multiple imaging feature encoders (DWI, MTT, and TTP), clinical metadata encoder, image and clinical metadata fusion b. 3D-based convolutional kernels, Inception V3 architecture	n.a.	5-fold cross validation

mRS, Modified Rankin Scale; HI, hemorrhagic infarction; HU, Hounsfield Units; PH, parenchymatous hematoma; DWI, diffusion-weighted imaging; MTT, mean transit time; ROI: regions of interest; TTP, time to peak; TICI, thrombolysis in cerebral infarction score; n.a., not available.

10.3389/fneur.2022.945813

TABLE 3 Model performance of conventional machine learning algorithms.

Clinical outcome	Imaging modality	Clinical variable	Model	References	Sample size (T/EV)	Model performance		Validation	
						AUC (95% CI)	Others	Internal	External
Good functional outcome at 90 days (mRS≤2 or mRS≤3)	NCCT and CTA	Yes	Gradient boosting decision trees	Brugnara et al. (22)	246	0.74 (0.73–0.75)	ACC,0.71	Yes	No
		Yes	RFA	Van et al. (23)	1,383*	0.79 (0.79–0.79)	n.a.	Yes	No
	NCCT	Yes	Regression trees	Alawieh et al. (24)	110/36	Internal: 0.93 (0.85–1.00)† External: n.a.	Internal: n.a. External: PV+: 0.60, NV-: 0.95	Yes	Yes
		Yes	RLR	Nishi et al. (25)	387/115	Internal: 0.86 (0.78–0.94)† External:0.90 (0.83–0.97)†	Internal: ACC,0.75; SEN,0.59; SPE,0.86; External: n.a.	Yes	Yes
	MRI (DWI and PWI)	Yes	RFA	Hamann et al. (26)	222	0.68 (0.61–0.76)	n.a.	Yes	No
		Yes	RFA	Kerleroux et al. (27)	133	0.83 (0.74–0.92)†	ACC,0.73; SEN,0.69; SPE,0.76	Yes	No
	MRI(DWI)	Yes	SVM	Xie et al. (28)	143	0.82 (0.75–0.89)†	ACC,0.77	Yes	No
Poor functional outcome at 90 days (mRS≥5 or mRS≥4)	NCCT and CTA	Yes	ANN	Ramos et al. (29)	1,401*	0.81 (0.79–0.83)	ACC, 0.65; SEN, 0.53; SPE,0.89; PV+, 0.69; NV-,0.80	Yes	No
	CTA	Yes	SVM	Ryu et al. (30)	482 (hold-out testing: 208)	0.82 (0.76–0.87)	n.a.	Yes	No
	n.a.	Yes	Decision trees	Kappelhof et al. (31)	1,090*	n.a	ACC,0.72	Yes	No
Successful reperfusion (TICI score≥2b)	NCCT and CTA	Yes	RFA	Van et al. (23)	1,383*	0.55 (0.55–0.56)	n.a.	Yes	No
Successful reperfusion at the first attempt (TICI score≥2b)	NCCT and CTA	No	RLR	Patel et al. (32)	119	0.77 (0.54–0.90)	ACC, 0.74	Yes	No
	NCCT and CTA	No	SVM	Hofmeister et al. (33)	109/47	External: 0.88 (0.75–1.00)†	External: ACC, 0.85; SEN, 0.50; SPE, 0.97, PV+, 0.86; NV-,0.85	Yes	Yes

ANN, artificial neural networks; AUC, area under the Receiver Operating Characteristic curve; ACC, accuracy; CTA, computed tomography angiography; DWI, diffusion weighted imaging; EV, external validation dataset; MRI, magnetic resonance imaging; NCCT, non-contrast computed tomography; NV-, negative predictive value; PWI, perfusion weighted imaging; PV+, positive predictive value; RFA, random forest analysis; RLR, regularized logistic regression; SVM, support vector machine; SEN, sensitivity; SPE, specificity; T, training dataset; n.a., not available/not applicable. Note: *model derived from patients registered in MR CLEAN Registry (38). †95% CI was estimated based on normal distribution.

10.3389/fneur.2022.945813

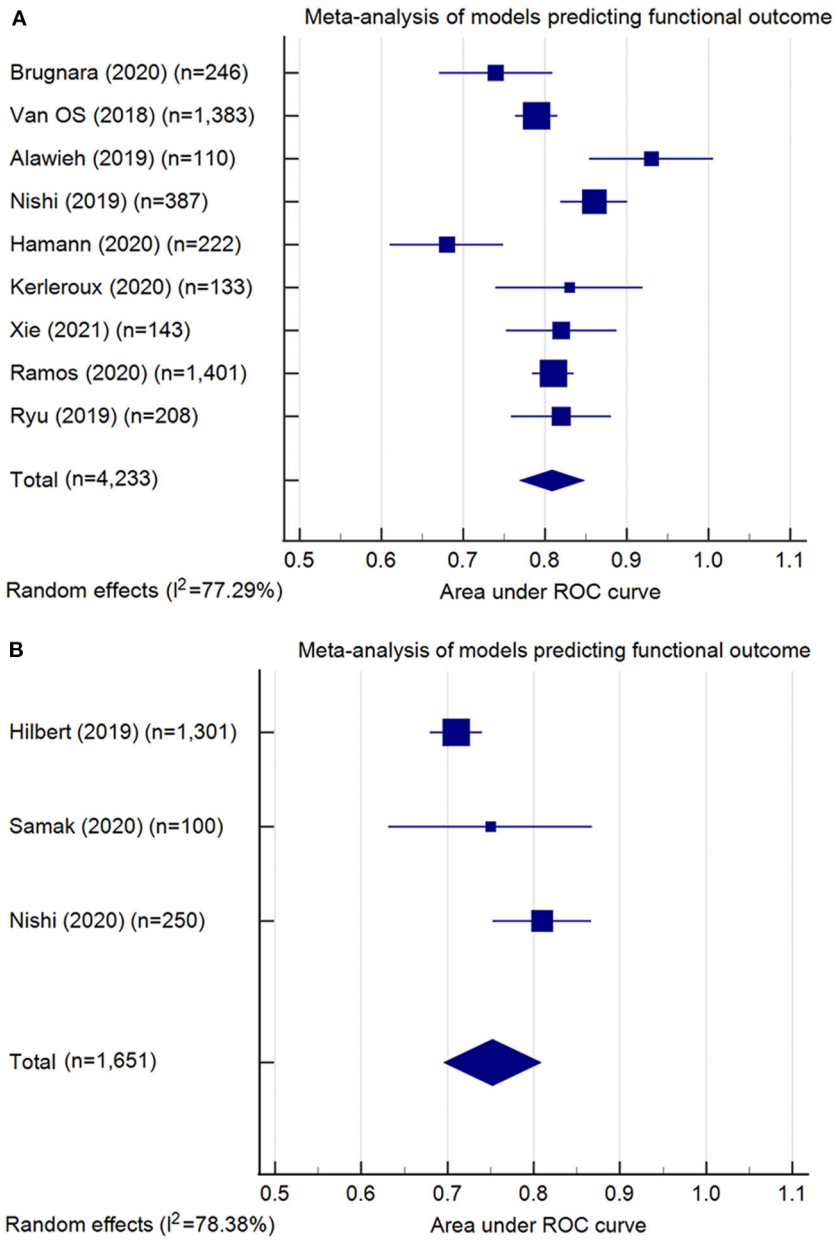


FIGURE 2
Meta-analysis of the area under the receiver-operating characteristics (ROC) curves (AUC) of models predicting functional outcome: **(A)** conventional machine learning models (pooled AUC = 0.81, 95% confidence interval: 0.77–0.85); **(B)** deep learning models (pooled AUC = 0.75, 95% confidence interval: 0.70–0.81). Note: Meta-analysis did not include the model developed by Kappelhof et al. (31), as the AUC was not reported.

a biased assessment of model performance if different thresholds were tested multiple times to obtain the "best" performance (19). Two studies (24, 35) also predicted each mRS point without dichotomizing the score, which may address a broader spectrum of functional status. On the other hand, a key outcome of clinical interest that remains un-investigated is futile recanalization, defined as poor functional outcomes at 90 days despite successful recanalization after EVT (42). Identification of those at high risk of futile recanalization is clinically and economically important, as an accurate prediction of this outcome would help avoid needless treatment and contribute to better resource allocation (42).

Accurate prediction of surrogate short-term outcomes may also help balance risk and benefit, and guide treatment approaches. There are two short-term outcomes investigated in the included studies—successful reperfusion and hemorrhagic

TABLE 4 Model performance of deep learning algorithms.

Clinical outcome	Imaging modality	Clinical variable	Model	References	Sample size (T/EV)	Model performance		Validation	
						AUC (95% CI)	Others	Internal	External
Good functional outcome at 90 days (mRS≤2)	CTA	No	DL (RFNN)	Hilbert et al. (34)	1,301	0.71(0.68–0.74)†	n.a.	Yes	No
	NCCT	Yes	DL (CNN)	Samak et al. (35)	400 (hold-out testing: 100)	0.75 (0.63–0.87)†	ACC,0.77	Yes	No
	MRI (DWI)	No	DL (CNN)	Nishi et al. (36)	250/74	Internal: 0.81 (0.70–0.92)† External:0.73 (0.61–0.85)†	Internal: SEN,0.76; SPE,0.76; ACC,0.72; External: SEN,0.72; SPE,0.60; ACC,0.65	Yes	Yes
Multiclass mRS (0, 1, 2, 3, 4, 5, 6) at 90 days	NCCT	Yes	DL (CNN)	Samak et al. (35)	400 (hold-out testing: 100)	n.a.	ACC, 0.35	Yes	No
Successful reperfusion (TICI score≥2b)	CTA	No	DL (RFNN)	Hilbert et al. (34)	1,301	0.65 (0.62–0.68)†	n.a.	Yes	No
Haemorrhagic transformation (including HI1, HI2, PH1, and PH2)	MRI (DWI and PWI)	Yes	DL (CNN)	Jiang et al. (37)	338/54	Internal: 0.95 (0.87–1.00)† External:0.94 (0.85–1.00)†	Internal: SEN, 0.86; SPE, 0.90; ACC,0.89; External: SEN,0.86; SPE,0.89; ACC,0.88	Yes	Yes

ANN, artificial neural networks; AUC, area under the Receiver Operating Characteristic curve; ACC, accuracy; CTA, computed tomography angiography; CNN, convolutional neural network; DWI, diffusion weighted imaging; EV, external validation dataset; HI, hemorrhagic infarction; PH, parenchymatous hematoma; MRI, magnetic resonance imaging; NCCT, non-contrast computed tomography; NV-, negative predictive value; PWI, perfusion weighted imaging; PV+, positive predictive value; RFA, random forest analysis; RLR, regularized logistic regression; RFNN, receptive field neural networks; SVM, support vector machine; SEN, sensitivity; SPE, specificity; T, training dataset; n.a., not available/not applicable. Note: *model derived from patients registered in MR CLEAN Registry (38). †95% CI was estimated based on normal distribution.

10.3389/fneur.2022.945813

transformation (HT). The model predicting HT (37) labeled all classes of HT as one category. However, it did not differentiate the symptomatic HT classes (i.e., PH2) from those classes without substantive mass effect (i.e., HT1 and HT2), and therefore may be of limited clinical utility. Also, there remains a gap in other relevant early outcomes. For example, occlusion at 24 h post-stroke, due to persistently failed recanalization or re-occlusion, has shown to be a predictor of longer-term outcomes in LVO patients (43) and may warrant investigation.

Missing data

Missing data has been a general problem in medical datasets and was the most common potential cause of bias in the reviewed studies. Potential bias may be introduced when data are missing conditional on the observed data (44), so a systematic approach to dealing with missing data will improve the quality of a study, and hence should be considered. For a ML model, data may be missing in outcomes (labels), covariates, and medical images. For the former two, there is substantial knowledge regarding how to deal with missing data (45). Multiple imputation is generally recommended, as it leads to minimum bias by imputing missing values while preserving the original data characteristics (19, 44). In terms of missing imaging data, there are currently no generally accepted mitigatory methods, although this is an area of active methodological research (46).

Model performance and limitations

Although conventional ML models can utilize a large quantity of clinical information, they have so far not demonstrated significant advantages against pre-treatment prognostic scores in predicting LVO outcomes (prognostic scores, AUC range: 0.61–0.80) (4). In contrast, a larger number of variables required in these models may limit the flexibility of their application in different clinical settings. Several conventional ML models achieved high performance values (AUCs: 0.86–0.93) (24, 25), but they were developed and validated in small datasets (sample size: development, 109–387; validation, 36–115) drawn from similar sampling frames (i.e., patients recruited in the same hospital at different time periods). ML models developed using small samples tend to be unstable and are likely to demonstrate substantially degraded predictive performance when applied to independent clinical populations (47). Overall, conventional ML models did not exhibit significant superiority when compared with prognostic scores for LVO outcome predictions.

Unlike conventional ML models that require variable selection, DL models are capable of analyzing raw imaging data in a "hypothesis-free" framework (9). However, the DL models in this review did not show superior performance to prognostic scores either. Most of these models were developed using small datasets, which may fail to capture the diverse features required to develop an accurate prognosis prediction model (48). This may also be one possible reason for the underwhelming performance. A few training schemes that suit clinical logics may help mitigate this issue (9). For example, augmenting data by mirroring CT images and inputting mirrored images with non-mirrored images enables the comparison between the affected side and the contralateral normal side, providing added information for model learning. Transfer learning from a clinically relevant task could also be a useful training scheme, e.g., pre-training the main task on an auxiliary task such as predicting occlusion of the left or right hemisphere. Further, multimodal data with richer information allows a model to capture diverse features and therefore may augment model performance. For example, multiple imaging modalities can provide diverse information, such as spatial information of hyperdense arteries, abnormal gray-white matter differentiation region and collateral supply (49). Similarly, non-imaging data can provide clinical-pathological features (i.e., blood glucose) that are associated with infarct progression and poor stroke outcomes (50). However, using multimodal imaging requires more computational resources, which may be a limiting factor for some research groups. Moreover, leveraging expert clinical knowledge is important to help augment model performance. For example, segmentation of hyperdense arteries or lesion and penumbra regions by experts allows additional information to be utilized in model development so that models can be trained to learn not only global features (i.e., location) but also fine details of the abnormal regions (i.e., boundary and shape).

Barriers to real-world implementation

There are several barriers currently that may impede the clinical utilization of the models described in the current review. Firstly, only five models (26.3%) reviewed were validated externally. External validation in an out-of-distribution population tests the robustness and stability of model performance across different populations. For example, model performance may be impacted when the imaging data for model development have certain characteristics derived from different scanners and image acquisition protocols. Indeed, a study focusing on predicting retinopathy showed that the model performance degraded significantly when images were taken under poor lighting conditions and with lower imaging resolution (51). External validation can help verify that model performance is not impacted by unexpected factors and can identify models that are more generalizable to diverse populations of LVO patients—this is critical information for implementation in a local clinical setting (21, 52). Conversely, it is not sufficient to demonstrate performance without external

validation (including prospective external validation) in a similar patient cohort. Over time, there are likely to be shifts in demographic composition and disease characteristics, as well as changes in new types of imaging scanners and image acquisition methods, even in the same center where that model was developed. A model tested only on an internal dataset may be brittle to these kinds of changes and see a drop in performance when used clinically. Secondly, only one study (32) published sufficient details of the models, including hyperparameters, coefficients (weights) and model equations, and only three studies (23, 26, 29) made the codes available online. Without the publication of sufficient details for independent model validation, it is difficult to directly implement published machine learning models in either validation studies or pre-clinical evaluation in local clinical environments. Current guidelines recommend the publication of "sufficient" details for validation, such as model structure, components, and values that used to control the learning process (hyperparameters) with code (19, 20, 53). For a deep learning model, it is difficult to publish millions of internal parameters in the paper, while it could be valuable to save files containing these parameters for future tasks as pre-trained weights. Thirdly, DL algorithms are usually described as "black box," which may limit their explainability and acceptability for patients, clinicians, and policymakers (54). Visualization techniques such as saliency maps (55, 56) are used to aid in model interpretability in two included studies (34, 36) and do so by highlighting the regions of an image that contribute most to a classification decision. However, these techniques themselves require cautious interpretation as they can highlight portions of an image with both clinically relevant and irrelevant information, and an image can still be misclassified based on such information (54). Explainability techniques are prone to offer false reassurance that a model is behaving in an appropriate manner, and we should instead depend on thorough performance evaluation to engender trust in DL systems (54).

Limitations and strengths

There are several limitations of our review. Firstly, we only included studies looking at LVO ischemic stroke treated with EVT treatment and did not examine studies including EVT for distal occlusion. However, as EVT is currently not a proven treatment for distal occlusion, any assessment of outcome prediction in this cohort is premature. Secondly, we have utilized re-calculated CIs for model comparisons and meta-analysis to ensure the similarity of the methods used. While most of the re-calculated 95% CIs are close to the original 95% CIs reported by included studies, we did note a significant deviation in the 95% CIs of two models (22, 32). We believe it is reasonable to rely on our wider estimate of CI compared to that provided in Brugnara et al. (22), as this original CI was extremely narrow based on a bootstrapping method and was much narrower than other 95% CIs reported on similar sized datasets. For the re-calculated CI that was narrower than the original report in Patel et al. (32), we again feel that the re-calculated version is more comparable to other studies as the small sample size resulted in less than 15 patients for the validation set, likely exaggerating the variability across cross-validation samples. This is the first comprehensive systematic review of ML and DL studies designed to predict clinical outcomes in LVO patients following EVT. Strengths of this review include a comprehensive literature search, independent screening and data extraction, as well as detailed quality assessment, all following PRISMA guidelines. More importantly, we conducted meta-analyses to quantitatively synthesize model performance, which has not been done in previous research that focused on ML and/or DL models for stroke prognostic prediction.

Conclusions

ML and DL algorithms have been evolving rapidly and are being increasingly applied to prognostic prediction of LVO patients treated with EVT. However, the application of ML and DL to this field is at an early stage. The outcomes investigated so far are limited, and further studies may consider additional clinically important outcomes, such as futile recanalization and post-treatment complications. High risk of potential bias due to missing data and lack of reporting details of prediction models were seen in most studies. Following PROBAST and TRIPOD guidelines can help improve study quality and reporting transparency. The performance of conventional ML and DL models did not substantially differ from each other or from the performance of pre-existing simple prognostic scores. Although a few ML models achieved high performance, most were developed using small datasets and lacked solid external validation. There is potential for ML outcome prediction techniques to be superior to conventional techniques, though larger/diverse datasets, more rigorous data preprocessing, and solid external validation, are required before incorporation into clinical practice.

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

MZ contributed to study conception, design, collection and analysis of data, and draft writing. LJP contributed to the study design, data collection, data analysis, and critical revision of the

manuscript. LOR contributed to the study design, data analysis, and critical revision of the manuscript. AB, LS, RS, TK, JJ, and MJ contributed to the study design and critical revision of the manuscript. ZW contributed to data collection and critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

MZ is supported by the Australian Government Research Training Program Scholarship. AB and LS are supported by the GlaxoSmithKline and the Australian Government Research Training Program Scholarship. LOR and LJP are supported by the GlaxoSmithKline. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

We would like to thank Drs. Stephan Lau, Gabriel Maicas, and Mr. Robert Franchini for their advice regarding strategies for literature searches.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2022.945813/full#supplementary-material

References

1. Malhotra K, Gornbein J, Saver JL. Ischemic strokes due to large-vessel occlusions contribute disproportionately to stroke-related dependence and death: a review. *Front Neurol.* (2017) 8:651. doi: 10.3389/fneur.2017.00651

2. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, et al. 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* (2018) 49:e46–e110. doi: 10.1161/STR.0000000000000158

3. Goyal M, Menon B, van Zwam W, Dippel DJ, Mitchell P, Demchuk A, 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

4. Kremers F, Venema E, Duvekot M, Yo L, Bokkers R, Lycklama ANG, et al. Outcome prediction models for endovascular treatment of ischemic stroke: systematic review and external validation. *Stroke.* (2021) 2021:STROKEAHA120033445. doi: 10.1161/STROKEAHA.120.033445

5. Kobkitsuksakul C, Tritanon O, Suraratdecha V. Interobserver agreement between senior radiology resident, neuroradiology fellow, and experienced neuroradiologist in the rating of alberta stroke program early computed tomography score (ASPECTS). *Diagn Interv Radiol.* (2018) 24:104–7. doi: 10.5152/dir.2018.17336

6. Nicholson P, Hilditch CA, Neuhaus A, Seyedsaadat SM, Benson JC, Mark I, et al. Per-region interobserver agreement of alberta stroke program early CT scores (ASPECTS). *J Neurointerv Surg.* (2020) 12:1069–71. doi: 10.1136/neurintsurg-2019-015473

7. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med.* (2016) 375:1216–9. doi: 10.1056/NEJMp1606181

8. Chauhan NK, Singh K, editors. A Review on Conventional Machine Learning vs Deep Learning. In: *2018 International Conference on Computing, Power and Communication Technologies (GUCON)* Greater Noida: IEEE (2018). doi: 10.1109/GUCON.2018.8675097

9. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005

10. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J.* (2015) 13:8–17. doi: 10.1016/j.csbj.2014.11.005

11. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J.* (2017) 38:500–7. doi: 10.1093/eurheartj/ehw188

12. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. *The Lancet.* (2011) 377:108–9. doi: 10.1016/S0140-6736(10)60903-8

13. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* (2009) 6:e1000097. doi: 10.1371/journal.pmed.1000097

14. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition.* (1997) 30:1145–59. doi: 10.1016/S0031-3203(96)00142-2

15. Hackshaw A. Statistical formulae for calculating some 95% confidence intervals. In: *A Concise Guide to Clinical Trials.* West Sussex: John Wiley & Sons, Ltd (2009). p. 205–7. doi: 10.1002/9781444311723.oth2

16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj.* (2003) 327:557–60. doi: 10.1136/bmj.327.7414.557

17. Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol.* (2000) 53:477–84. doi: 10.1016/S0895-4356(99)00204-8

18. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021).*). Cochrane Database of Systematic Reviews (2022). Available online at: www.training.cochrane.org/handbook (accessed August 26, 2022).

19. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med.* (2019) 170:51–8. doi: 10.7326/M18-1376

20. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg.* (2015) 102:148–58. doi: 10.1002/bjs.9736

21. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ.* (2020) 368:m689. doi: 10.1136/bmj.m689

22. Brugnara G, Neuberger U, Mahmutoglu MA, Foltyn M, Herweh C, Nagel S, et al. Multimodal predictive modeling of endovascular treatment outcome for acute ischemic stroke using machine-learning. *Stroke.* (2020) 2020:3541–51. doi: 10.1161/STROKEAHA.120.030287

23. 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

24. Alawieh A, Zaraket F, Alawieh MB, Chatterjee AR, Spiotta A. Using machine learning to optimize selection of elderly patients for endovascular thrombectomy. *J Neurointerv Surg.* (2019) 11:847–51. doi: 10.1136/neurintsurg-2018-014381

25. 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

26. Hamann J, Herzog L, Wehrli C, Dobrocky T, Bink A, Piccirelli M, et al. Machine learning based outcome prediction in stroke patients with MCA-M1 occlusions and early thrombectomy. *Eur J Neurol.* (2020) 21:14651. doi: 10.1111/ene.14651

27. Kerleroux B, Benzakoun J, Janot K, Dargazanli C, Eraya DD, Ben Hassen W, et al. Relevance of brain regions' eloquence assessment in patients with a large ischemic core treated with mechanical thrombectomy. *Neurology.* (2021) 97:e1975–e85. doi: 10.1212/WNL.0000000000012863

28. Xie Y, Oster J, Micard E, Chen B, Douros IK, Liao L, et al. Impact of pretreatment ischemic location on functional outcome after thrombectomy. *Diagnostics (Basel).* (2021) 11:2038. doi: 10.3390/diagnostics11112038

29. Ramos LA, Kappelhof M, van Os HJA, Chalos V, Van Kranendonk K, Kruyt ND, et al. Predicting poor outcome before endovascular treatment in patients with acute ischemic stroke. *Front Neurol.* (2020) 11:580957. doi: 10.3389/fneur.2020.580957

30. Ryu CW, Kim BM, Kim HG, Heo JH, Nam HS, Kim DJ, et al. Optimizing outcome prediction scores in patients undergoing endovascular thrombectomy for large vessel occlusions using collateral grade on computed tomography angiography. *Clin Neurosurg.* (2019) 85:350–8. doi: 10.1093/neuros/nyy316

31. Kappelhof N, Ramos LA, Kappelhof M, van Os HJA, Chalos V, van Kranendonk KR, et al. Evolutionary algorithms and decision trees for predicting poor outcome after endovascular treatment for acute ischemic stroke. *Comput Biol Med.* (2021) 133:104414. doi: 10.1016/j.compbiomed.2021.104414

32. Patel TR, Waqas M, Sarayi S, Ren Z, Borlongan CV, Dossani R, et al. Revascularization outcome prediction for a direct aspiration-first pass technique (ADAPT) from pre-treatment imaging and machine learning. *Brain Sci.* (2021) 11:1321. doi: 10.3390/brainsci11101321

33. Hofmeister J, Bernava G, Rosi A, Vargas MI, Carrera E, Montet X, et al. Clot-based radiomics predict a mechanical thrombectomy strategy for successful recanalization in acute ischemic stroke. *Stroke.* (2020) 51:2488–94. doi: 10.1161/STROKEAHA.120.030334

34. 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

35. Samak ZA, Clatworthy P, Mirmehdi M. Prediction of thrombectomy functional outcomes using multimodal data. Medical image understanding and analysis. *Commun Comput Inf Sci.* (2020) 2020:267–79. doi: 10.1007/978-3-030-52791-4_21

36. Nishi H, Oishi N, Ishii A, Ono I, Ogura T, Sunohara T, et al. Deep learning-derived high-level neuroimaging features predict clinical outcomes for large vessel occlusion. *Stroke.* (2020) 51:1484–92. doi: 10.1161/STROKEAHA.119.028101

37. Jiang L, Zhou L, Yong W, Cui J, Geng W, Chen H, et al. A deep learning-based model for prediction of hemorrhagic transformation after stroke. *Brain Pathol.* (2021) 2021:e13023. doi: 10.1111/bpa.13023

38. Jansen IGH, Mulder M, Goldhoorn RB. Endovascular treatment for acute ischaemic stroke in routine clinical practice: prospective, observational cohort study (MR CLEAN Registry). *BMJ.* (2018) 360:k949. doi: 10.1136/bmj.k949

39. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke.* (1988) 19:604–7. doi: 10.1161/01.STR.19.5.604

40. Bhat SS, Fernandes TT, Poojar P, da Silva Ferreira M, Rao PC, Hanumantharaju MC, et al. Low-field MRI of stroke: challenges and opportunities. *J Magn Reson Imaging.* (2021) 54:372–90. doi: 10.1002/jmri.27324

41. Wintermark M, Reichhart M, Cuisenaire O, Maeder P, Thiran JP, Schnyder P, et al. Comparison of admission perfusion computed tomography and qualitative diffusion- and perfusion-weighted magnetic resonance imaging in acute stroke patients. *Stroke.* (2002) 33:2025–31. doi: 10.1161/01.STR.0000023579.61630.AC

42. Nie X, Pu Y, Zhang Z, Liu X, Duan W, Liu L. Futile recanalization after endovascular therapy in acute ischemic stroke. *Biomed Res Int.* (2018) 2018:5879548. doi: 10.1155/2018/5879548

43. Marto JP, Strambo D, Hajdu SD, Eskandari A, Nannoni S, Sirimarco G, et al. Twenty-four-hour reocclusion after successful mechanical thrombectomy: associated factors and long-term prognosis. *Stroke.* (2019) 50:2960–3. doi: 10.1161/STROKEAHA.119.026228

44. Thomas RM, Bruin W, Zhutovsky P, van Wingen G. Chapter 14 - dealing with missing data, small sample sizes, and heterogeneity in machine learning studies of brain disorders. In: Mechelli A, Vieira S, editors. *Machine Learning.* London: Academic Press (2020). p. 249–66. doi: 10.1016/B978-0-12-815739-8.00014-6

45. Little RJ, Rubin DB. *Statistical Analysis With Missing Data, Vol. 793.* In: Little RJ, Rubin DB, editors. Hoboken, NJ: John Wiley & Sons (2019). doi: 10.1002/9781119013563

46. Yoon J, Jordon J, Schaar M, editors. Gain: Missing Data Imputation Using Generative Adversarial Nets. In: *International Conference on Machine Learning.* Vienna: PMLR (2018). p. 5689–98. doi: 10.48550/arXiv.1806.02920

47. Cui Z, Gong G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage.* (2018) 178:622–37. doi: 10.1016/j.neuroimage.2018.06.001

48. Sun C, Shrivastava A, Singh S, Gupta A, editors. Revisiting unreasonable effectiveness of data in deep learning era. In: *Proceedings of the IEEE International Conference on Computer Vision.* Venice: IEEE (2017).

49. Merino JG, Warach S. Imaging of acute stroke. *Nat Rev Neurol.* (2010) 6:560–71. doi: 10.1038/nrneurol.2010.129

50. Lindsberg PJ, Roine RO. Hyperglycemia in acute stroke. *Stroke.* (2004) 35:363–4. doi: 10.1161/01.STR.0000115297.92132.84

51. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Honolulu, HI: Association for Computing Machinery (2020). p. 1–12.

52. Geirhos R, Jacobsen J-H, Michaelis C, Zemel R, Brendel W, Bethge M, et al. Shortcut learning in deep neural networks. *Nat Mach Intell.* (2020) 2:665–73. doi: 10.1038/s42256-020-00257-z

53. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force−7. *Value Health.* (2012) 15:843–50. doi: 10.1016/j.jval.2012.04.012

54. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digital Health.* (2021) 3:e745–e50. doi: 10.1016/S2589-7500(21)00208-9

55. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision.* Venice: IEEE (2017).

56. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems.* Long Beach, CA (2017). p. 4768–77. doi: 10.48550/arXiv.1705.07874

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 12 October 2022
DOI 10.3389/fneur.2022.1034402

OPEN ACCESS

EDITED BY
Baofeng Gao,
Beijing Institute of Technology, China

REVIEWED BY
Yang Wang,
Capital Medical University, China
Yingkun He,
Henan Provincial People's Hospital, China

*CORRESPONDENCE
Wei Sun
sunwei7777@126.com
Youxiang Li
liyouxiang@bjtth.org

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

SPECIALTY SECTION
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

RECEIVED 01 September 2022
ACCEPTED 27 September 2022
PUBLISHED 12 October 2022

CITATION
Gao H, You W, Wei D, Lv J, Sun W and Li Y (2022) Tortuosity of parent artery predicts in-stent stenosis after pipeline flow-diverter stenting for internal carotid artery aneurysms. *Front. Neurol.* 13:1034402. doi: 10.3389/fneur.2022.1034402

Tortuosity of parent artery predicts in-stent stenosis after pipeline flow-diverter stenting for internal carotid artery aneurysms

Haibin Gao[1,2,3†], Wei You[1,4†], Dachao Wei[1,4], Jian Lv[1,4], Wei Sun[2,3*] and Youxiang Li[1,4*]

[1]Department of Interventional Neuroradiology, Beijing Neurosurgical Institute and Beijing Tiantan Hospital, Capital Medical University, Beijing, China, [2]Department of Neurosurgery, Beijing Boai Hospital, China Rehabilitation Research Center, Beijing, China, [3]College of Rehabilitation, Capital Medical University, Beijing, China, [4]Beijing Engineering Research Center, Beijing, China

Background and purpose: The relationship between the tortuosity of the parent artery and treatment outcomes is not well established. We investigate the association between parent artery tortuosity and flow diverter (FD) treatment outcomes in patients with internal carotid artery aneurysms in this study.

Methods: A retrospective review study was conducted to identify all patients with internal carotid artery aneurysms who were implanted with Pipeline embolization device (PED) between 2016 and 2020. The relationship between parent artery tortuosity and aneurysm complete occlusion (CO) and in-stent stenosis (ISS) was analyzed. The mathematical parameters "Curvature", "torsion", and "DM" extracted from the parent artery were utilized to quantify the parent artery tortuosity. A vascular narrowing of greater than 25% was categorized as ISS. Logistic regression analysis was used to identify significant independent predictors. Furthermore, we compared the performance of four machine learning algorithms and Logistic Regression model in predicting ISS.

Results: This research included 62 patients who with internal carotid artery aneurysms. In 49 (79%) cases, follow-up angiography (mean follow-up duration 11.7 ±7.3 months) revealed CO of the aneurysm. ISS was detected in 22 (35.5%) cases. According to univariate analysis, parent artery tortuosity and other variables were not associated with CO ($p > 0.1$). Maximum curvature (OR = 1.084; 95% CI, 1.008–1.165; $p = 0.03$) and DM (OR = 0.01; 95% CI, 0–0.488; $p = 0.02$) exhibited strong independent associations with ISS in multivariate analysis. The SVM model is superior to the conventional Logistic Regression model and the other models in predicting ISS.

Conclusions: The tortuosity of the parent artery may affect the treatment outcome of FD stenting. We found that parent artery tortuosity was associated with ISS, but not with aneurysm complete occlusion following PED stenting

for internal carotid artery aneurysms in this study. Parent arteries with higher maximum curvature and lower DM were more likely to develop ISS.

KEYWORDS

Intracranial Aneurysms, Pipeline embolization device (PED), vascular tortuosity, in stent stenosis, internal carotid artery (ICA)

Introduction

Having gained widespread global acceptance, flow diverters (FD) have ushered in a paradigm shift in the treatment of IAs (1). The Pipeline embolization device (PED) is one of the earliest and most widely used FD and was initially approved for the treatment of internal carotid artery aneurysms (ICA) (2). A total of 83.6% of PED stents were used to treat internal carotid artery aneurysms, mostly located in the carotid siphon and carotid supraclinoid (3). Despite its short length, this segment of the artery exhibits complex morphology and marked population variation.

Tortuosity of vessels is a common angiographic finding that can be associated with vascular pathologies and may suggest systemic diseases, such as hypertension or diabetes mellitus (4–6). In terms of intracranial vasculature, tortuosity is associated with aneurysms, Moyamoya disease, and the presence of atherosclerosis (7–9). Furthermore, vascular tortuosity is associated with hemodynamic changes and vessel wall remodeling in coronary and peripheral artery studies (10, 11).

Hemodynamic changes are the main mechanism of PED stent treatment of aneurysms. ISS is a wellknown but understudied consequence after endovascular stents implantation, and was assumed to be associated with an inflammatory response due to hemodynamic status and intimal injury (12, 13). Little is known about the relationship between the degree of tortuosity of the parent artery and treatment outcomes after the PED stenting for the aneurysm. This study aimed to investigate the correlation between parent artery tortuosity and aneurysm complete occlusion (CO) and in-stent stenosis following Pipeline Flow-Diverter Stenting for Intracranial Aneurysms.

Methods

Study population

We retrospectively reviewed the consecutive patients with internal carotid artery aneurysms who received PED treatment at the Interventional Neuroradiology Department of our hospital from 2016 to 2020. The inclusion criteria for the study population were as follows: (1) Patients with sufficient quality of pre-stenting 3D rotational angiography (3DRA) imaging for the 3D artery model reconstruction and the analysis of tortuosity without incomplete and missing image sequence, non-standardized protocol, and severe motion artifact. (2) Patients with at least one digital subtraction angiography (DSA) follow-up for angiographic evaluations. (3) Patients who had successfully received PED implantation, while the target aneurysm without received any previous stenting and coiling before visiting our hospital. (4) Patients presenting with ideal results of parent artery reconstruction evaluated by neurosurgeons.

A total of 226 consecutive patients with internal carotid artery aneurysm who were treated with a PED stent and underwent at least one digital subtraction angiography (DSA) follow-up were retrospectively reviewed in the present study. After excluding patients without the adequate quality of pre-stenting 3D rotational angiography (3DRA) imaging (n = 143), 7 patients who received the previous stenting and coiling (n = 7) and those without idealized 3D model reconstruction of the parent artery (n = 14) were enrolled in this study. Patient demographics, aneurysm characteristics, procedural information, and clinical and angiographic outcomes were reviewed. This retrospective study was approved, and patient's written consent was waived off by our institutional review board.

Endovascular procedure

The patients' treatment was started by receiving dual antiplatelet medication with aspirin 100 mg/day and clopidogrel 75 mg/day for 7 days before the implantation. Routine preoperative platelet function tests were performed, and the patients who were identified as clopidogrel non-responders received either prasugrel or ticagrelor. All PED implantations were performed under general anesthesia through a femoral approach. According to the aneurysm anatomy and based on the operator's experience, the treatment strategy was formulated to decide whether PED was to be used alone or with coiling. For patients with incomplete release of stent or incomplete stent apposition, the stent was massaged using a wire or with a balloon angioplasty. After the procedure, dual antiplatelet therapy was maintained for 6 months, and aspirin was continued indefinitely thereafter.

Assessment of aneurysm CO and ISS

Aneurysm occlusion was determined using follow-up digital subtraction angiography, and grading was established based on two angiographic views using the O'Kelly-Marotta grading scale (14). Follow-up DSA images were referred to assess the loss rate of the parent artery diameter, as shown in the DSA images, as the gap between the lumen vessel was filled with contrast material and the stent strut. ISS was defined as vessel narrowing of >25%. The loss rate of the parent artery diameter was calculated as follows: 1 – (vessel diameter/stent diameter) × 100% (Figure 1).

The measurement of ISS was performed by neuroradiologists with at least 3 years of experience and then reviewed by a senior neuroradiologist.

Assessment of artery tortuosity

The tortuosity analysis was based on the vessel centerline that was considered to represent the main geometric attribute of the vessel. In this study, we selected the stent implantation segment of the parent artery based on pre-stenting 3DRA imaging for artery tortuosity analyses. The mathematical parameters "Curvature", "torsion", and "DM" were extracted from the center line and then used to quantitatively evaluate the parent artery tortuosity (Figure 2).

Idealized 3D model of the parent artery

Based on the raw 3D acquisition data, the 3D model was extracted by segmentation and surface reconstruction tools based on thresholding in Mimics 19.0 (Materialize, Belgium). Smoothing operations and removal of the unrelated branching were then conducted on the 3D surface, and the domain inlets and outlets were truncated perpendicular to the centerline using the Geomagic Studio 2012 (North Carolina). The VMTK software was applied to automatically remove the aneurysm on the 3D model so as to achieve the reconstruction of the parent artery. The reconstructed 3D model was then evaluated by 2 senior interventional neuroradiologists with >5 years of experience for the successful reconstruction. Cases wherein the two neuroradiologists evaluated the results of reconstruction inconsistently were excluded from this study. The stent implantation segment of the parent artery was then identified and segmented based on the proximal and distal positions of the implanted PED stent.

Centerline extraction and tortuosity metrics calculation

The centerline was defined as the locus of the centers of the maximal-inscribed spheres along the vessel itself (15). The centerlines of the reconstructed 3D model were calculated automatically in the Aneufuse software as a set of discrete 3D points, which was used as an input to obtain an analytical representation through 3D freeknots regression splines (16, 17). The curvature and torsion of each discrete 3D point and the distance metric (DM) of the parent artery were calculated using a customized working program. The curvature of a curve at a point was defined geometrically as the inverse of the radius of the osculating circle at that particular point. Torsion was defined as a measure of how sharply a curve twisted out of the plane of the curvature (18). The DM quantifies the "lengthening effect" of tortuosity and was calculated as follows: $DM = l/L$, where l is the straight-line distance from the beginning point to the end point of the segment and L is the total path length of the centerline. For each parent artery, the mean curvature, maximum curvature, and range curvature were calculated. Using the same logic, the values of mean torsion, maximum torsion, and range torsion were calculated.

Establishment of machine learning model

A total of 75 variables were included in the model establishment. Continuous variables were standardized with z-score transformation. In order to avoid class imbalance, the Boardline SMOTE algorithm was applied to the dataset. After preprocessing, the dataset was randomly split into training set (80%) and test set (20%). We used Recursive Feature Elimination (RFE) to select the best combination of ISS predictors. Then traditional Logistic Regression and four machine learning algorithms (elastic net [ENT], support vector machine [SVM], Xgboost [XGB], and random forest [RF]) were developed to predict the occurrence of ISS with the open-source machine learning library scikit-learn (version 0.24.1). Then in model training, 10-fold cross-validation and grid research were used to determine the optimal hyperparameters of the models. We compared accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) in the test set to find the best prediction model. Finally, we use the Shapley additive explanation (SHAP) algorithm (version 0.39.0) to calculate the feature importance.

Statistical analysis

Data were presented as the frequency for categorical variables and as the mean with range for continuous variables. The Chi-square test or Fisher's exact test was applied to analyze the categorical variables, and the independent samples t-test was applied to analyze the continuous variables. Univariate and multivariate analyses were used to analyze the relationship between tortuosity of the parent artery and the occlusion of aneurysm and in-sent stenosis. Binary logistic regression

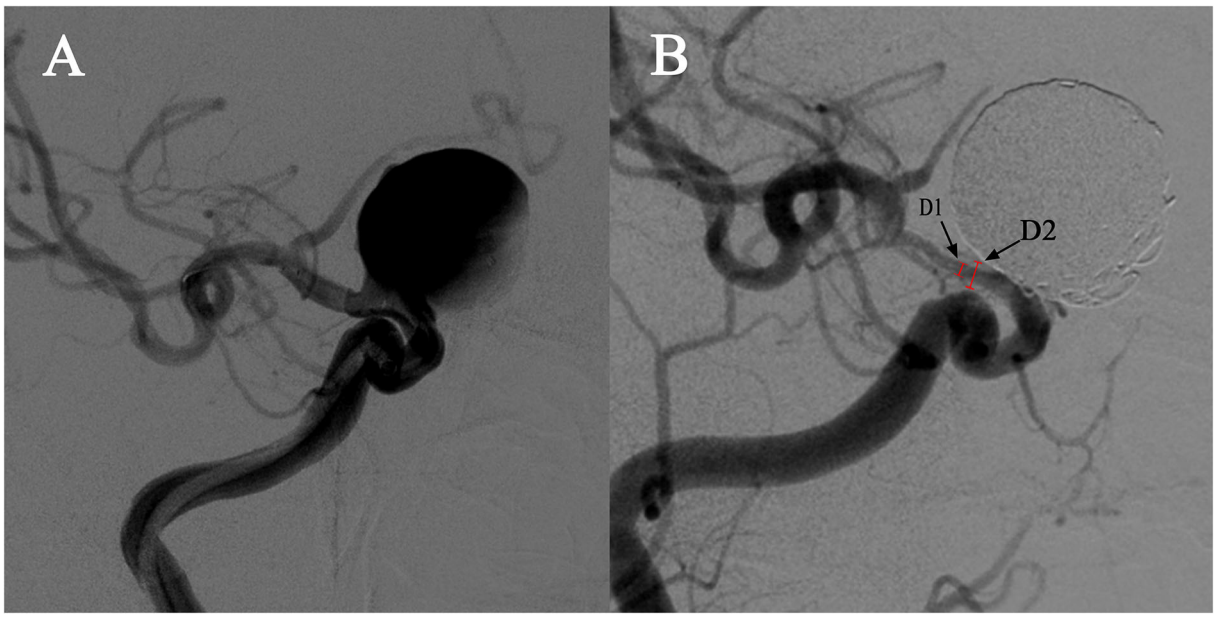


FIGURE 1
Demonstration of quantitative determination of stent stenosis. **(A)** A case of a right carotid ophthalmic aneurysm treated with a PED stent and coiling. **(B)**. The follow-up angiography showed in-stent stenosis at the distal end of the stent. D1 is the narrowest vessel diameter and D2 is the stent diameter in the same position confirmed by the mask image. The stenosis was calculated as follows: 1 − [D2/D1] × 100%.

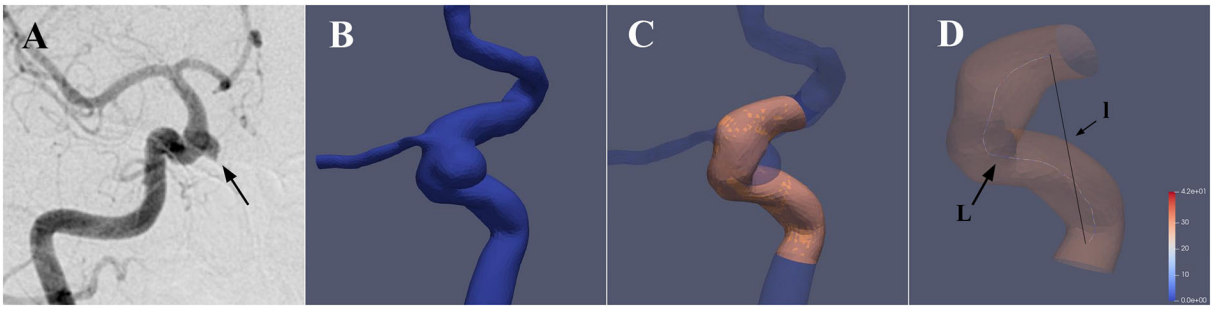


FIGURE 2
3D reconstruction of the parent artery and extraction of the vessel centerline. **(A)** A case with a right carotid ophthalmic aneurysm was treated with a PED stent. **(B)** The 3D vessel model was extracted from the raw 3D acquisition data using segmentation and surface reconstruction tools. **(C)** The aneurysm and irrelevant branch vessels were removed to get the idealized 3D reconstruction model of the parent artery, and the stent implantation segment was then identified and segmented. **(D)** Aneufuse software was used to determine the centerlines of the reconstructed 3D model. The curvature and torsion of each discrete 3D point and distance metric (DM) of the parent artery were then calculated. (l is the straight-line distance and L is the total path length)

analysis was used to identify significant independent predictors. Variables that were found to be significant at the level of 0.1 under univariate analysis or based on clinical relevance were subjected to binary logistic regression analysis. The results are presented in the form of an odds ratio (OR) and a corresponding 95% confidence interval (CI). $p < 0.05$ was considered to indicate statistical significance. A receiver operating characteristic (ROC) curve was used to analyze the performance of the logistic regression classification model. Accordingly, we performed the statistical analysis and plotted the figures using SPSS and GraphPad software.

Results

Patient demographics, aneurysm characteristics, procedure details, and clinical outcomes

Table 1 shows the demographics, aneurysm characteristics, and angiographic outcomes of the patients. The PED stent was used on 62 patients (mean age of 54.2 ±9.2 years; 47 females, 75.8%) with 62 targeted internal carotid artery aneurysms. All cases were saccular side-wall aneurysms, with the majority

TABLE 1 Univariate analysis in association with CO of aneurysm.

Variables	All (n = 62)	nCO (n = 13)	CO (n = 49)	p
Baseline demographics and clinical characteristics				
Female, no. (%)	47 (75.8)	9 (69.2)	38 (77.6)	0.716
Age, y (mean ± SD)	54.2±9.2	52.5±9.1	54.7±9.3	0.468
BMI	25.1±3.7	26.5±4.7	24.8±3.3	0.13
Co-morbidity				
Hypertension, no. (%)	27 (43.5)	6 (46.2)	21 (42.9)	1
Diabetes, no. (%)	5 (8.1)	0 (0)	5 (10.2)	0.574
Hyperlipidemia, no. (%)	24 (38.7)	3 (23.1)	21 (42.89)	0.222
History of allergies, no. (%)	7 (11.3)	1 (7.7)	6 (12.2)	1
Smoking, no. (%)	13 (21.0)	4 (30.8)	9 (18.4)	0.444
Alcohol abuse, no. (%)	9 (14.5)	3 (23.1)	6 (12.2)	0.381
Symptomatic presentation of IA, no. (%)	36 (58.1)	10 (76.9)	26 (53.1)	0.205
Ruptured (history of SAH), no. (%)	3 (4.8)	2 (15.4)	1 (2)	0.109
Aneurysm characteristics				
Aneurysm neck size (mm)	6.2±2.8	6±2.6	6.3±2.9	0.705
Maximum diameter (mm)	9.4±4.7	9.3±3.7	9.4±4.9	0.928
Parent artery diameter (mm)	3.8±0.7	3.7±1	3.8±0.6	0.73
Associate with parent artery stenosis, no. (%)	3 (4.8)	1 (7.7)	2 (4.1)	0.513
Procedure characteristics				
PED plus coiling, no. (%)	30 (48.4)	4 (30.8)	26 (53.1)	0.215
PED Flex, no. (%)	37 (59.7)	10 (76.9)	27 (55.1)	0.21
Multiple PED implantation, no. (%)	8 (12.9)	3 (23.1)	5 (10.2)	0.347
Balloon angioplasty, no. (%)	14 (22.6)	4 (30.8)	10 (20.4)	0.466
Tortuous parameters of parent artery				
Mean curvature	0.6±0.5	0.5±0.3	0.6±0.5	0.564
Maximum curvature	5.3±8.9	4.1±4.8	5.7±9.7	0.579
Range curvature	5.3±8.9	4.1±4.8	5.6±9.7	0.579
Mean torsion	12.4±4.2	11.8±4.1	12.6±4.3	0.563
Maximum torsion	45.4±16.2	43.4±15.3	45.9±16.5	0.625
Range torsion	0.1±0.2	43.4±15.3	45.8±16.6	0.635
DM	0.5±0.2	0.5±0.2	0.5±0.2	0.823
L (total path length)	23.9±9.5	25.5±9.2	23.5±9.7	0.507
l (straight line distance)	12.1±6.2	13.5±8.3	11.8±5.5	0.368

of the aneurysms found in the C6 segment (39/62, 62.9%), 11 (17.7%) in the C7, 6 (9.7%) in the C5, and 6 (9.7%) in the C4. The mean aneurysm maximum length and neck size were 9.4 ±4.7 mm and 6.2 ±2.8 mm, respectively. Ruptured aneurysms accounted for 4.8% (3/62) of cases. There were 32 (51.6%) cases treated with PED alone and 30 (48.4%) cases treated with PED plus coiling. PED Flex was used in 37 (59.7%) cases, whereas PED classic was used in the remaining cases. Multiple PED implantation was used in 8 (12.9%) procedures, while balloon angioplasty was administered in 14 (22.6%) procedures. The mean procedure duration was 121.2 ±59.5 min.

In 49 (79%) cases, follow-up angiography (mean follow-up duration 11.7 ±7.3 months) revealed CO of aneurysms. ISS was detected in 22 (35.5%) lesions with a mean follow-up time of 7.8 ±4.4 months. ISS with > 50% vessel narrowing in 3 cases. There were no symptomatic cases of ISS. Treatment-related complications were observed in 3 (4.8%) cases during the periprocedural period, with one case of aneurysm rupture during the treatment procedure, one case of parenchymal hemorrhage, and one case of infarction occurring after the treatment procedure (< 24 h). There were two-thirds of cases (mRS<2) with transient deficits and none with permanent deficits.

TABLE 2 Univariate and multivariate logistic analysis in association with ISS.

Variables	ISS (n = 40)	ISS (n = 22)	Univariate	Multivariate	
			p	p	OR (95%CI)
Baseline demographics and clinical characteristics					
Female, no. (%)	31 (77.5)	16 (72.7)	0.76		
Age, *y* (mean ± SD)	54.1±8.8	54.4±10.2	0.923		
BMI	25.1±3.8	25.2±3.6	0.906		
Co-morbidity					
Hypertension, no. (%)	14 (35)	13 (59.1)	0.067	0.065	—
Diabetes, no. (%)	2 (5)	3 (13.6)	0.337		
Hyperlipidemia, no. (%)	18 (45)	6 (27.3)	0.188		
History of allergies, no. (%)	4 (10)	3 (13.6)	0.989		
Smoking, no. (%)	9 (22.5)	4 (18.2)	0.756		
Alcohol abuse, no. (%)	6 (15)	3 (13.6)	1		
Symptomatic presentation of IA, no. (%)	20 (50)	16 (72.7)	0.109		
Ruptured (history of SAH), no. (%)	2 (5)	1 (4.5)	1		
Aneurysm characteristics					
Aneurysm neck size (mm)	6.4±2.8	5.9±2.8	0.567		
Maximum diameter (mm)	9.9±4.7	8.5±4.6	0.275		
Parent artery diameter (mm)	3.8±0.6	3.7±0.8	0.594		
Associate with parent artery stenosis, no. (%)	1 (2.5)	2 (9.1)	0.59		
Procedure characteristics					
PED plus coiling, no. (%)	22 (55)	8 (36.4)	0.192		
PED Flex, no. (%)	27 (67.5)	10 (45.5)	0.09	0.288	—
Multiple PED implantation, no. (%)	4 (10)	4 (18.2)	0.438		
Balloon angioplasty, no. (%)	9 (22.5)	5 (22.7)	1		
Tortuous parameters of parent artery					
Mean curvature	0.5±0.3	0.8±0.7	0.054	0.549	—
Maximum curvature	3.4±6	8.8±11.9	0.021	0.03	1.084 (1.008–1.165)
Range curvature	3.4±6	8.8±11.9	0.021		
Mean torsion	11.8±3.7	13.5±4.9	0.138		
Maximum torsion	43.8±16.5	48.2±15.6	0.312		
Range torsion	43.7±16.5	48.2±15.6	0.308		
DM	0.6±0.2	0.5±0.1	0.021	0.02	0.01 (0–0.488)
L (total path length)	23.7±10.8	24.3±6.9	0.814		
l (straight line distance)	12.9±7.1	10.7±3.8	0.12		

Assessment of artery tortuosity

Table 1 summarizes the findings of the examination of parent artery tortuosity. The mean, maximum, and range curvatures of the parent arteries were 0.6 ±0.5, 5.3 ±8.9, and 5.3 ±8.9, respectively. Parent artery mean torsion, maximum torsion, and range torsion were 12.4 ±4.2, 45.4 ±16.2, and 45.3 ±16.2, respectively. The "L" and "l" of parent arteries were 23.9 ±9.5 and 12.1 ±6.2 mm, respectively, and the DM was 0.5 ±0.2.

Parent artery tortuosity with CO and ISS

The relationship between parent artery tortuosity and the CO of aneurysm and ISS was analyzed (Tables 1, 2). Univariate analysis showed that parent artery tortuosity and other variables were not associated with CO ($p > 0.1$). The maximum curvature of the parent artery was significantly higher in individuals with ISS than in those without ISS (8.8 ±11.9 vs. 3.4 ±6; $p = 0.021$). The DM of the parent artery with ISS was significantly smaller than those without ISS (0.5 ±0.1 vs. 0.6 ±0.2; $p = 0.021$).

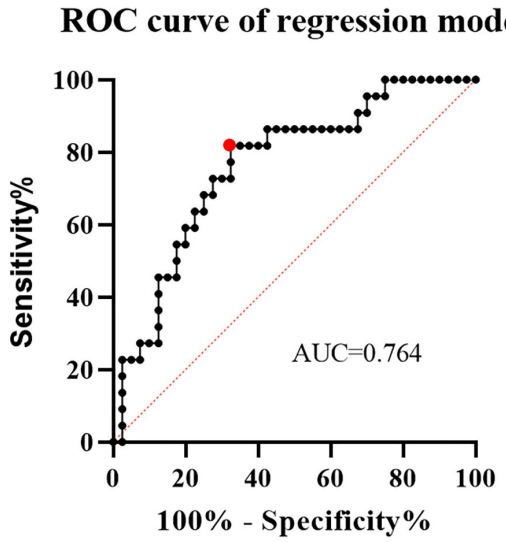


FIGURE 3
The receiver operating characteristic (ROC) curve of a logistic regression classification model. At the cutoff value, the area under the curve (AUC) was 0.764, with sensitivity and specificity of 0.82 and 0.68, respectively.

Baseline demographics and clinical characteristics, aneurysm characteristics, and procedure characteristics did not differ significantly between patients with ISS and without ISS ($p > 0.1$). Significant variables in the univariate analysis that met the threshold of 10% were subjected to multivariate regression. After variables with collinearity were excluded from the collinearity test findings, mean curvature, maximum curvature, DM, and hypertension were included in the multivariate regression. In the multivariate analysis, the maximum curvature (OR = 1.084; 95% CI, 1.008–1.165; $p = 0.03$) and DM (OR = 0.01; 95% CI, 0–0.488; $p = 0.02$) exhibited strong independent associations with ISS. Specifically, parent arteries with higher maximum curvature and smaller DM were more likely to develop ISS. ROC curve was used to analyze the performance of the logistic regression classification model and the area under the curve (AUC) was 0.764 (Figure 3).

ISS prediction using machine learning

In order to explore the predictive value of parent artery tortuosity on ISS, we built machine learning models. After RFE, 6 predictors (height, DM, maximum curvature, aneurysm neck, hypertension and dyslipidemia) were identified. In the training set, the RF model had the highest mean AUC-ROC (0.956; 95% confidence interval [CI], 0.951–0.961), followed by the XGB model (0.951; 95% CI, 0.943–0.959), the SVM model (0.883; 95% CI, 0.874–0.893), the ENT model (0.793; 95% CI, 0.775–0.811) and the LR model (0.789; 95% CI, 0.770–0.808) (Figure 4A). In the validation set, the SVM had the best mean AUC-ROC (0.762; 95% CI, 0.626–0.899), followed by the ENT model (0.725; 95% CI, 0.562–0.888), the LR model (0.721; 95% CI, 0.563–0.878), the XGB model (0.661; 95% CI, 0.461–0.860), the RF model (0.659; 95% CI, 0.514–0.805) (Figure 4B). In the test set, the SVM model had the best performance (0.891), with an accuracy of 87.5%, a sensitivity of 100% and a specificity of 75%, followed by the XGB model (0.875), the RF model (0.859), the ENT model (0.797) and the LR model (0.734) (Figure 4C). In SHAP analysis, DM had the greatest impact on prediction model, followed by height, hypertension, aneurysm neck, maximum curvature, and dyslipidemia (Figure 4D). Height, hypertension, and maximum curvature are the risk factors of ISS, while DM, aneurysm neck and dyslipidemia are the protective factors of ISS.

Discussion

In the present study, we investigated the correlation between vessel tortuosity of the parent artery, in-stent stenosis, and aneurysm complete occlusion following PED Stenting for internal carotid artery aneurysm. To the best of our knowledge, this is the first study to analyze the impact of parent artery tortuosity on FD treatment results using a quantitative measurement method. "Curvature" is a parameter describing geometric curvature, "torsion" is a parameter describing the degree to which the curve torsion deviates from the osculating plane, and DM is a parameter describing the degree to which the curve deviates from the straight line. In general, more tortuous vessels have higher curvature, higher torsion, and lower DM. We discovered that vessel tortuosity was associated with ISS, but not with CO in our research. The multivariable analysis determined that higher maximum curvature and lower DM were independent predictors of in-stent stenosis. The SVM model is superior to the conventional Logistic Regression model and the other models in predicting ISS.

Previous literature has reported significant disparities in the incidence of ISS after PED implantation, which may be due to different definitions of ISS (18, 19). Some authors (20–22) considered vessel narrowing of <25% as intimal hyperplasia or vessel narrowing of >25% as in-stent stenosis, which was consistent with our study's definition of ISS. ISS is a wellknown issue of endovascular stent implantation, although the underlying cellular mechanisms of ISS have not been welldescribed. We hypothesize that ISS formation is related to complex relationships between hemodynamics, vascular biology, and mechanical properties of the parent vessel and the stent. Previous research has revealed that abnormal vascular remodeling and neointimal hyperplasia are potential causes of ISS (23). Endothelial dysfunction caused by vascular endothelial injury leads to the proliferation of local smooth muscle cells, the formation of neointimal tissue, and eventually ISS (13).

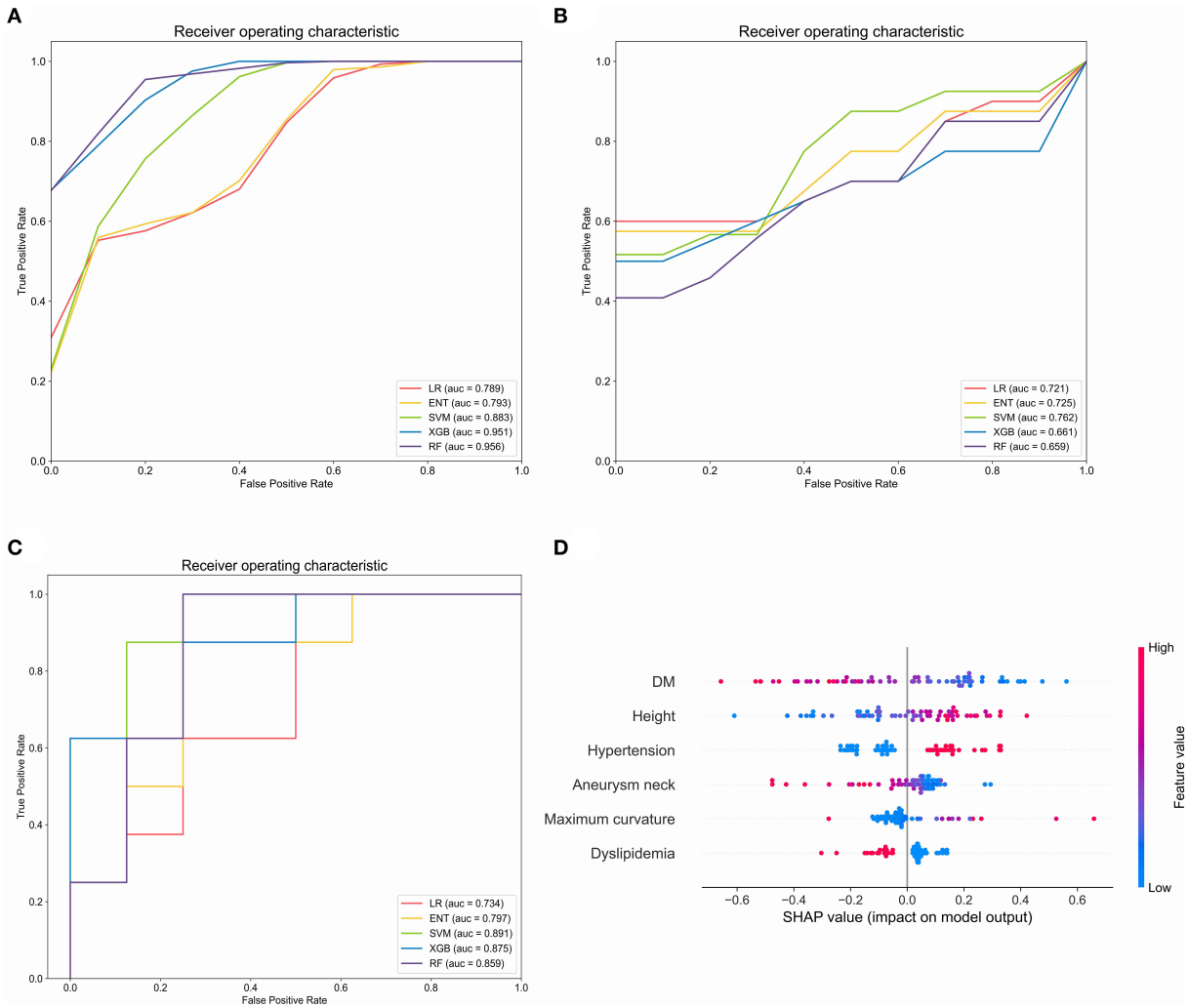


FIGURE 4
Evaluation of the machine learning models in the training set, validation set and test set. **(A)** Comparison of mean area under the receiver operating curve (AUC-ROC) in five prediction models in the training set. **(B)** Comparison of AUC-ROC in five prediction models in the validation set. **(C)** Comparison of AUC-ROC in five prediction models in the test set. **(D)** Feature importance (mean |SHAP value|) of each predictor. LR, Logistic regression; ENT Elastic net regression; Xgb, Xgboost; RF, random forest. The red or blue color of the points represents high or low feature value and the left or right of the X-axis represents positive or negative influence on the model, respectively.

The arterial wall is invariably damaged during endovascular procedures, causing local inflammation and smooth muscle cell proliferation, leading to intimal hyperplasia and stenosis. Compared to straight arteries, stent implantation may cause greater damage to the tortuous region, resulting in fibroblastic and neointimal hyperplasia (24), leading to ISS. Additionally, the forces between the stent and the vessel wall are more complex and unstable in curved vessels than in straight vessels, which may lead to more severe vascular endothelial injury and inflammatory response.

Coronary and peripheral artery studies have shown that vascular tortuosity can affect arterial hemodynamics. Increased arterial tortuosity decreases perfusion pressure, wall shear stress, and prolonged relative residence time (25). Lowering wall shear stress can contribute to matrix metalloproteinase activation (26), leading to arterial wall remodeling (27). According to hemodynamic simulations, a curved bend disturbs the constant flow characteristic of straight vessels, resulting in complex and heterogeneous flow patterns (28, 29), which may reduce arterial endothelial function and trigger pathological degeneration of the arterial wall, which favors ISS formation (30–32).

In order to further explore the predictive value of parent artery tortuosity on ISS, we developed ISS prediction models.

In contrast with traditional LR model, machine learning models can solve non-linear problem and multicollinearity, which may improve the prediction performance of the model. Therefore, we compared the performance of four popular machine learning models and LR model. The results indicate that the SVM model is superior to the conventional Logistic Regression model and the other models in terms of AUC-ROC, accuracy, sensitivity, and specificity. In SHAP analysis, DM is the most important predictor of ISS. Height, hypertension, and maximum curvature are the risk factors of ISS, while aneurysm neck and dyslipidemia are protective factors of ISS, which is consistent with the results of multivariable analysis.

The effect of FD is dependent on the induced hemodynamic changes that trigger the process of thrombosis and endothelial remodeling and ultimately seal the aneurysm. Szikora et al. (33) pointed out that the angle between the aneurysm and the parent vessel was the most important determinant of blood flow pattern in the sac. Furthermore, Xu et al. (34) showed that the hemodynamic changes of aneurysms after FD implantation were strongly dependent on the curvature of the parent artery. They found that as the curvature of the parent artery increased, the pressure, inflow velocity, and inflow volume rate also increased, while the aneurysm sac's relative residence time decreased. However, we did not find an association between parent artery tortuosity and aneurysm occlusion in this study, but this does not prove that the morphology of the parent artery does not affect the outcome of aneurysm occlusion. The arterial tortuosity parameters in this study reflect the overall tortuosity evaluation of the parent artery and may be incapable of assessing the tortuosity characteristics of specific regions, which might explain the unfavorable results.

Various methods have been proposed for the analysis of vascular tortuosity in two-dimensional (2D) and three-dimensional (3D) (10, 35). Lang and Reiter examined 89 head halves in 1984 and determined three patterns of carotid siphon morphology based on the curve angle of the artery bend (36). In 1965, Weibel and Fields proposed a classification method of internal carotid artery morphology based on the degree of vessel angulation in 2D angiography images (37). They classified the internal carotid artery vascular tortuosity patterns as Kinking, Looping, and Coiling. Many studies have examined vascular tortuosity using mathematical metrics in addition to the aforementioned assessment methods based on global vessel morphology (9, 38). To quantify vascular tortuosity, they employed parameters including the sum of angle metrics, product of angle distance, triangular index, and inflection count metrics obtained from 2D angiographic images. To evaluate vascular tortuosity, recent studies have used quantitative mathematical parameters such as tortuosity and tortuosity based on vascular reconstruction images (18, 39). In this study, 3D analysis was employed since it is more accurate and has higher consistency and accuracy than 2D analysis.

This research has certain drawbacks. First, since this is a retrospective study, it does not imply that ISS and parent artery tortuosity are causally related. To investigate the mechanism of their association, further prospective studies and laboratory evidence are needed. Second, this was a single-center retrospective study that only included patients who underwent 3D rotational angiography, which may have limited the study sample size and increased the risk of selection bias. The relatively small number of study sample may also weaken the generalization of machine learning models. Third, this study focused on the association between global vessel tortuosity and treatment outcomes, and future studies should focus on better assessing vessel tortuosity in specific regions. Despite these limitations, this is the first study to analyze the FD treatment results on parent artery tortuosity using a quantitative 3D analysis method.

Conclusion

The tortuosity of the parent artery may affect the treatment outcome of FD stenting. We discovered that parent artery tortuosity was associated with ISS, but not with complete aneurysm occlusion after PED stenting for internal carotid artery aneurysms in this study. ISS was more common in parent arteries with higher maximum curvature and lower DM. To corroborate the current study's findings, larger cohort prospective studies and a more comprehensive assessment of vascular tortuosity are needed.

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.

Author contributions

YL, WY, and HG designed the study. WY, JL, and DW contributed to data collection and data analysis. WY and HG drafted the manuscript. WS and HG performed the revision of the current literature. All authors contributed to the manuscript and approved.

Funding

This study has received funding by the National Natural Science Foundation of China (82171289), the

National Key Research and Development Program of China (2017YFB1304400).

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

References

1. Gory B, Berge J, Bonafe A, Pierot L, Spelle L, Piotin M, Biondi A, Cognard C, Mounayer C, Sourour N, Barbier C, Desal H, Herbreteau D, Chabert E, Brunel H, Ricolfi F, Anxionnat R, Decullier E, Huot L, Turjman F, Investigators D. Flow diverters for intracranial aneurysms: the diversion national prospective cohort study. *Stroke.* (2019) 50:3471–80. doi: 10.1161/STROKEAHA.119.024722

2. Ito K, Kai Y, Hyodo A, Ishiuchi S. Long-term outcome of angioplasty or stent placement for stenosis of the cavernous or petrous portion of the internal carotid artery. *Neurol Med Chir.* (2011) 51:813–8. doi: 10.2176/nmc.51.813

3. Luo B, Kang H, Zhang H, Li T, Liu J, Song D, Zhao Y, Guan S, Maimaitili A, Wang Y, Feng W, Wang Y, Wan J, Mao G, Shi H, Yang X. Pipeline embolization device for intracranial aneurysms in a large Chinese cohort: factors related to aneurysm occlusion. *Ther Adv Neurol Disord.* (2020) 13:1756286420967828. doi: 10.1177/1756286420967828

4. Owen CG, Newsom RS, Rudnicka AR, Barman SA, Woodward EG, Ellis TJ. Diabetes and the tortuosity of vessels of the bulbar conjunctiva. *Ophthalmology.* (2008) 115:e27–32. doi: 10.1016/j.ophtha.2008.02.009

5. Li Y, Shen C, Ji Y, Feng Y, Ma G, Liu N. Clinical implication of coronary tortuosity in patients with coronary artery disease. *PLoS ONE.* (2011) 6:e24232. doi: 10.1371/journal.pone.0024232

6. Sasongko MB, Wong TY, Nguyen TT, Cheung CY, Shaw JE, Wang JJ. Retinal vascular tortuosity in persons with diabetes and diabetic retinopathy. *Diabetologia.* (2011) 54:2409–16. doi: 10.1007/s00125-011-2200-y

7. Ruan LT, Duan YY, Cao TS, Zhuang L, Huang L. Color and power doppler sonography of extracranial and intracranial arteries in Moyamoya disease. *J Clin Ultrasound.* (2006) 34:60–9. doi: 10.1002/jcu.20201

8. Kim BJ, Kim SM, Kang DW, Kwon SU, Suh DC, Kim JS. Vascular tortuosity may be related to intracranial artery atherosclerosis. *Int J Stroke.* (2015) 10:1081–6. doi: 10.1111/ijs.12525

9. Klis KM, Krzyzewski RM, Kwinta BM, Stachura K, Gasowski J. Tortuosity of the internal carotid artery and its clinical significance in the development of aneurysms. *J Clin Med.* (2019) 8:237. doi: 10.3390/jcm8020237

10. Ciurica S, Lopez-Sublet M, Loeys BL, Radhouani I, Natarajan N, Vikkula M, Maas A, Adlam D, Persu A. Arterial tortuosity. *Hypertension.* (2019) 73:951–60. doi: 10.1161/HYPERTENSIONAHA.118.11647

11. Hughes AD, Martinez-Perez E, Jabbar AS, Hassan A, Witt NW, Mistry PD, Chapman N, Stanton AV, Beevers G, Pedrinelli R, Parker KH, Thom SA. Quantification of topological changes in retinal vascular architecture in essential and malignant hypertension. *J Hypertens.* (2006) 24:889–94. doi: 10.1097/01.hjh.0000222759.61735.98

12. Aguilar Perez M, Bhogal P, Henkes E, Ganslandt O, Bazner H, Henkes H. In-stent stenosis after p64 flow diverter treatment. *Clin Neuroradiol.* (2018) 28:563–8. doi: 10.1007/s00062-017-0591-y

13. Kipshidze N, Dangas G, Tsapenko M, Moses J, Leon MB, Kutryk M, Serruys P. Role of the endothelium in modulating neointimal formation: vasculoprotective approaches to attenuate restenosis after percutaneous coronary interventions. *J Am Coll Cardiol.* (2004) 44:733–9. doi: 10.1016/S0735-1097(04)01083-6

14. O'Kelly JC, Krings T, Fiorella D, Marotta TR. A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents. *Interv Neuroradiol J Peritherapeutic Neuroradiol Surg Proced Related Neurosci.* (2010) 16:133–7. doi: 10.1177/159101991001600204

15. Archie Jr JP. Geometric dimension changes with carotid endarterectomy reconstruction. *J Vasc Surg.* (1997) 25:488–98. doi: 10.1016/S0741-5214(97)70259-3

16. Gallo D, Steinman DA, Morbiducci U. An insight into the mechanistic role of the common carotid artery on the hemodynamics at the carotid bifurcation. *Ann Biomed Eng.* (2015) 43:68–81. doi: 10.1007/s10439-014-1119-0

17. Gallo D, Vardoulis O, Monney P, Piccini D, Antiochos P, Schwitter J, Stergiopulos N, Morbiducci U. Cardiovascular morphometry with high-resolution 3D magnetic resonance: first application to left ventricle diastolic dysfunction. *Med Eng Phys.* (2017) 47:64–71. doi: 10.1016/j.medengphy.2017.03.011

18. Vorobtsova N, Chiastra C, Stremler MA, Sane DC, Migliavacca F, Vlachos P. Effects of vessel tortuosity on coronary hemodynamics: an idealized and patient-specific computational study. *Ann Biomed Eng.* (2016) 44:2228–39. doi: 10.1007/s10439-015-1492-3

19. Texakalidis P, Bekelis K, Atallah E, Tjoumakaris S, Rosenwasser RH, Jabbour P. Flow diversion with the pipeline embolization device for patients with intracranial aneurysms and antiplatelet therapy: a systematic literature review. *Clin Neurol Neurosurg.* (2017) 161:78–87. doi: 10.1016/j.clineuro.2017.08.003

20. John S, Bain MD, Hui FK, Hussain MS, Masaryk TJ, Rasmussen PA, Toth G. Long-term follow-up of in-stent stenosis after pipeline flow diversion treatment of intracranial aneurysms. *Neurosurgery.* (2016) 78:862–7. doi: 10.1227/NEU.0000000000001146

21. Deutschmann HA, Wehrschuetz M, Augustin M, Niederkorn K, Klein GE. Long-term follow-up after treatment of intracranial aneurysms with the Pipeline embolization device: results from a single center. *AJNR Am J Neuroradiol.* (2012) 33:481–6. doi: 10.3174/ajnr.A2790

22. Lylyk P, Miranda C, Ceratto R, Ferrario A, Scrivano E, Luna HR, Berez AL, Tran Q, Nelson PK, Fiorella D. Curative endovascular reconstruction of cerebral aneurysms with the pipeline embolization device: the Buenos Aires experience. *Neurosurgery.* (2009) 64:632–42. doi: 10.1227/01.NEU.0000339109.98070.65

23. Buccheri D, Piraino D, Andolina G, Cortese B. Understanding and managing in-stent restenosis: a review of clinical data, from pathogenesis to treatment. *J Thorac Dis.* (2016) 8:E1150–62. doi: 10.21037/jtd.2016.10.93

24. Albuquerque FC, Fiorella D, Han P, Spetzler RF, McDougall CG. A reappraisal of angioplasty and stenting for the treatment of vertebral origin stenosis. *Neurosurgery.* (2003) 53:607–14. doi: 10.1227/01.NEU.0000079494.87390.28

25. Rikhtegar F, Knight JA, Olgac U, Saur SC, Poulikakos D, Marshall Jr W, Cattin PC, Alkadhi H, Kurtcuoglu V. Choosing the optimal wall shear parameter for the prediction of plaque location—a patient-specific computational study in human left coronary arteries. *Atherosclerosis.* (2012) 221:432–7. doi: 10.1016/j.atherosclerosis.2012.01.018

26. Berceli SA, Jiang Z, Klingman NV, Pfahnl CL, Abouhamze ZS, Frase CD, Schultz GS, Ozaki CK. Differential expression and activity of matrix metalloproteinases during flow-modulated vein graft remodeling. *J Vasc Surg.* (2004) 39:1084–90. doi: 10.1016/j.jvs.2003.12.031

27. Ota R, Kurihara C, Tsou TL, Young WL, Yeghiazarians Y, Chang M, Mobashery S, Sakamoto A, Hashimoto T. Roles of matrix metalloproteinases in flow-induced outward vascular remodeling. *J Cereb Blood Flow Metab.* (2009) 29:1547-58. doi: 10.1038/jcbfm.2009.77

28. Foutrakis GN, Yonas H, Sclabassi RJ. Saccular aneurysm formation in curved and bifurcating arteries. *AJNR Am J Neuroradiol.* (1999) 20:1309–17. doi: 10.1016/j.atherosclerosis.2010.03.001

29. Zhang C, Xie S, Li S, Pu F, Deng X, Fan Y, Li D. Flow patterns and wall shear stress distribution in human internal carotid arteries: the geometric effect on the risk for stenoses. *J Biomech*. (2012) 45:83–9. doi: 10.1016/j.jbiomech.2011.10.001

30. Sugiyama S, Niizuma K, Nakayama T, Shimizu H, Endo H, Inoue T, Fujimura M, Ohta M, Takahashi A, Tominaga T. Relative residence time prolongation in intracranial aneurysms: a possible association with atherosclerosis. *Neurosurgery*. (2013) 73:767–76. doi: 10.1227/NEU.0000000000000096

31. Riccardello Jr GJ, Shastri DN, Changa AR, Thomas KG, Roman M, Prestigiacomo CJ, Gandhi CD. Influence of relative residence time on side-wall aneurysm inception. *Neurosurgery*. (2018) 83:574–81. doi: 10.1093/neuros/nyx433

32. Krzyzewski RM, Klis KM, Kwinta BM, Gackowska M, Stachura K, Starowicz-Filip A, Thompson A, Gasowski J. Analysis of anterior cerebral artery tortuosity: association with anterior communicating artery aneurysm rupture. *World Neurosurg*. (2019) 122:e480–6. doi: 10.1016/j.wneu.2018.10.086

33. Szikora I, Paal G, Ugron A, Nasztanovics F, Marosfoi M, Berentei Z, Kulcsar Z, Lee W, Bojtar I, Nyary I. Impact of aneurysmal geometry on intraaneurysmal flow: a computerized flow simulation study. *Neuroradiology*. (2008) 50:411–21. doi: 10.1007/s00234-007-0350-x

34. Xu J, Wu Z, Yu Y, Lv N, Wang S, Karmonik C, Liu JM, Huang Q. Combined effects of flow diverting strategies and parent artery curvature on aneurysmal hemodynamics: a CFD study. *PLoS ONE*. (2015) 10:e0138648. doi: 10.1371/journal.pone.0138648

35. Bullitt E, Gerig G, Pizer SM, Lin W, Aylward SR. Measuring tortuosity of the intracerebral vasculature from MRA images. *IEEE Trans Med Imaging*. (2003) 22:1163–71. doi: 10.1109/TMI.2003.816964

36. Lang J, Reiter U. Course of the cranial nerves in the lateral wall of the cavernous sinus. *Neurochirurgia*. (1984) 27:93–7. doi: 10.1055/s-2008-1053667

37. Weibel J, Fields WS. Tortuosity, coiling, and kinking of the internal carotid artery. II. relationship of morphological variation to cerebrovascular insufficiency. *Neurology*. (1965) 15:462–8. doi: 10.1212/WNL.15.5.462

38. Kim BJ, Lee SH, Kwun BD, Kang HG, Hong KS, Kang DW, Kim JS, Kwon SU. Intracranial aneurysm is associated with high intracranial artery tortuosity. *World Neurosurg*. (2018) 112:e876-e880. doi: 10.1016/j.wneu.2018.01.196

39. Lauric A, Safain MG, Hippelheuser J, Malek AM. High curvature of the internal carotid artery is associated with the presence of intracranial aneurysms. *J Neurointerv Surgery*. (2014) 6:733–9. doi: 10.1136/neurintsurg-2013-010987

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 12 October 2022
DOI 10.3389/fneur.2022.1027557

OPEN ACCESS

EDITED BY
Baofeng Gao,
Beijing Institute of Technology, China

REVIEWED BY
Chao Fu,
China-Japan Union Hospital of Jilin University, China
Yingkun He,
Henan Provincial People's Hospital, China
Cong-Hui Li,
The First Hospital of Hebei Medical University, China

*CORRESPONDENCE
Chuanzhi Duan
doctor_duanzj163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Endovascular and Interventional Neurology, a section of the journal Frontiers in Neurology

RECEIVED 25 August 2022
ACCEPTED 27 September 2022
PUBLISHED 12 October 2022

CITATION
Tian Z, Li W, Feng X, Sun K and Duan C (2022) Prediction and analysis of periprocedural complications associated with endovascular treatment for unruptured intracranial aneurysms using machine learning. *Front. Neurol.* 13:1027557. doi: 10.3389/fneur.2022.1027557

Prediction and analysis of periprocedural complications associated with endovascular treatment for unruptured intracranial aneurysms using machine learning

Zhongbin Tian[1†], Wenqiang Li[2†], Xin Feng[1], Kaijian Sun[1] and Chuanzhi Duan[1*]

[1]National Key Clinical Specialty, Engineering Technology Research Center of Education Ministry of China, Guangdong Provincial Key Laboratory on Brain Function Repair and Regeneration, Neurosurgery Institute, Department of Neurosurgery, Zhujiang Hospital, Southern Medical University, Guangzhou, China, [2]Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Background: The management of unruptured intracranial aneurysm (UIA) remains controversial. Recently, machine learning has been widely applied in the field of medicine. This study developed predictive models using machine learning to investigate periprocedural complications associated with endovascular procedures for UIA.

Methods: We enrolled patients with solitary UIA who underwent endovascular procedures. Periprocedural complications were defined as neurological adverse events resulting from endovascular procedures. We incorporated three machine learning algorithms into our prediction models: artificial neural networks (ANN), random forest (RF), and logistic regression (LR). The Shapley Additive Explanations (SHAP) approach and feature importance analysis were used to identify and prioritize significant features associated with periprocedural complications.

Results: In total, 443 patients were included. Forty-eight (10.83%) procedure-related complications occurred. In the testing set, the ANN model produced the largest value (0.761) for area under the curve (AUC). The RF model also achieved an acceptable AUC value of 0.735, while the AUC value of the LR model was 0.668. SHAP and feature importance analysis identified distal aneurysm, aneurysm size and treatment modality as most significant features for the prediction of periprocedural complications following endovascular treatment for UIA.

Conclusion: Periprocedural complications after endovascular treatment for UIA are not negligible. Prediction of periprocedural complications *via* machine learning is feasible and effective. Machine learning can serve as a promising tool in the decision-making process for UIA treatment.

KEYWORDS

intracranial aneurysm, endovascular treatment, periprocedural complication, machine learning (ML), feature importance analysis

Introduction

The prevalence of unruptured intracranial aneurysm (UIA) in the adult population is about 3–7% (1, 2). The rupturing of UIAs usually results in subarachnoid hemorrhage, which is associated with a high rate of mortality and morbidity (3). In recent decades, endovascular treatment has become the first-line of treatment for intracranial aneurysm and has achieved satisfactory outcomes (4). However, most UIAs have a low annual risk of rupture, and complications related to the endovascular treatment of UIAs should not be neglected (5, 6). It remains controversial whether UIA should be treated or not. For these reasons, the risk of complications from endovascular treatment should be carefully weighed against the risk of UIA rupture. Establishing a method to identify factors associated with procedure-related complications, and to predict risk from such complications, could provide critical reference guidelines to physicians.

Recent studies have applied machine learning (ML) to the prediction of intracranial aneurysm rupture and outcome after endovascular treatment (7–9). When challenged with complex non-linear relationships across large datasets, ML can generate automated decisions that often outperform conventional statistical methods. Liu et al. and Zhu et al. reported promising results from the application of ML techniques to the prediction of aneurysm stability (10, 11). Paliwal et al. and Guédon et al. developed ML models to predict occlusion outcomes from aneurysms following flow diverter deployment (12, 13). However, research on the prediction of periprocedural complications from endovascular treatment is still scarce.

In this study, we exploited three ML algorithms to develop predictive models for periprocedural complications after endovascular treatment: artificial neural networks (ANN), random forest (RF), and logistic regression (LR). We then compared their prediction performance. To improve model interpretability and identify significant factors associated with periprocedural complications, we applied the Shapley Additive Explanations (SHAP) method and feature importance analysis. Our results provide physicians with reference guidelines for the management of UIAs.

Methods

Patient selection

This retrospective study was approved by the relevant institutional ethics committee, and written informed consent was obtained from patients or their relatives during hospitalization. We included patients with solitary unruptured saccular intracranial aneurysm that were treated endovascularly between January 2016 and December 2019. We adopted the following exclusion criteria: dissecting aneurysm, previous treatment, covered stent deployment, treatment performed by parent artery occlusion, and the existence of a brain arteriovenous malformation. On the basis of these criteria, we retained 443 cases for this study.

Endovascular procedures

The specific strategy for endovascular treatment was determined by a neurovascular team and was individually tailored to each case. Following general anesthesia, the endovascular procedure was performed. All patients received systemic intravenous heparin. If the team determined that it was necessary to deploy a conventional stent or a flow diverter, the endovascular procedure was preceded by a 5-day dual antiplatelet therapy (100 mg/d of aspirin and 75 mg/d of clopidogrel). If the team opted for a conventional stent, patients were advised to take clopidogrel (75 mg/d) for 6 weeks and aspirin (100 mg/d) for 6 months. If the flow diverter was deployed, the patient would take clopidogrel (75 mg/d) for 3 months and aspirin (100 mg/d) on a permanent basis thereafter.

Outcome measures

We recorded periprocedural complications that occurred within 30 days of the endovascular procedure. We divided the 443 cases into two groups: complication group and control group. Patients with periprocedural complications were assigned to the complication group. Periprocedural complications were defined as any neurological adverse event (increase in modified Rankin Scale score) resulting from the endovascular treatment. An adverse event was defined as major if the associated neurological deficit lasted longer than 7 days, otherwise it was defined as minor (14).

Clinical and morphological features

We analyzed the following factors: age, elderly status (>65 years of age), gender, potential risk factors (history of cigarette smoking and alcohol intake, hypertension, cardiovascular disease, hyperlipidemia, diabetes, and previous cerebral ischemic comorbidities), treatment modality (coiling only, stent-assisted coiling, or flow diverter treatment), aneurysm size (maximum size), presence of large aneurysm (size $\geq$ 10 mm), aneurysm neck size, presence of wide-neck aneurysms ($\geq$4 mm or dome-neck ratio $\leq$2), location (anterior/posterior circulation), shape (defined as irregular if presenting blebs, nipples, or multiple lobes), and presence of distal aneurysm (distally to the Circle of Willis).

Machine learning model development

We randomly divided data samples into training set (310 cases) and testing set (133 cases) with a 7:3 ratio. Because the dataset was imbalanced between complication and control groups, we applied an adaptive synthetic (ADASYN) sampling method to generate more synthetic data for the minority class (complication group) in the training set (15). After application of ADASYN, the training set was expanded to 553 cases (280 complication cases). We then trained three ML algorithms (ANN, RF, and LR) on the training set with ten-fold cross validation and grid search to optimize hyperparameters for each model. Details of the ML models are provided in Supplementary material. The testing set was used to estimate model performance. Model performance was evaluated *via* receiver operating characteristic (ROC) analysis. To improve model interpretability and investigate important features associated with perioperative complications, we used the SHAP method and feature importance analysis (16). We used the SHAP method to explore important features in ANN models. We used feature score/coefficient to evaluate feature importance in RF and LR models.

Statistical analysis

We performed statistical analyses using version 22.0 of SPSS (IBM Corp., Armonk, NY, USA). Data are presented as mean and standard deviation for quantitative variables, and as frequency for qualitative variables. We used univariate logistic analysis to analyze risk factors related to periprocedural complications after endovascular procedure for UIA. Statistical significance was defined as $p<0.05$.

Results

Patient and aneurysm characteristics

We enrolled a total of 443 patients (281 females and 162 males) for this study. Mean age was 55.97 ± 11.41 years. Mean aneurysm size was 6.92 ± 5.08 mm. Of the 443 cases, 75 cases were treated with coil embolization only, 270 with stent-assisted coiling, and 98 with flow diverter therapy.

Periprocedural complications

In total, 48 (10.83%) procedure-related complications occurred: 4 intraprocedural aneurysm ruptures (0.90%), 2 postprocedural aneurysm ruptures (0.45%), 2 cases of cranial nerve palsy (0.45%), and 40 ischemic events (9.03%). The 40 ischemic events included 26 ischemic strokes, 7 transient ischemic attacks, 4 intra-stent thrombosis and 3 thrombosis resulting from coil migration. Of these 48 procedure-related complications, 27 cases (6.09%) were associated with minor adverse events that resolved on discharge, and 21 (4.74%) were associated with major adverse events.

Risk factors for periprocedural complications

Table 1 shows results from univariate logistic regression analysis of risk factors for periprocedural complications. The age of the complication group was significantly older than that of the control group (59 ± 12 vs. 56 ± 11, $p = 0.043$), and the complication group included a higher proportion of elderly patients than the control group (35.4% vs. 16.2%, $p = 0.002$). Patients with hypertension and distal aneurysm showed a tendency toward more periprocedural complications ($p = 0.006$ and $p = 0.045$, respectively). Aneurysm size was significantly larger in the complication group than in the control group (8.67 ± 5.54 mm vs. 6.71 ± 4.98 mm, $p = 0.014$), and the complication group had larger aneurysms than the control group (31.3% vs. 16.7%, $p = 0.016$). Moreover, the incidence of periprocedural complications was higher in cases treated by flow diverter therapy (13.3%) or stent-assisted coiling (11.1%) than cases treated by coiling only (6.7%), although this difference did not reach statistical significance ($p = 0.167$ and $p = 0.265$, respectively).

Model performance and identification of important features

With relation to the training set, the area under the curve (AUC) value for the ANN model [0.993; 95% confidence interval (CI) 0.985–0.999] was similar to the AUC value for the RF model (0.999; 95% CI 0.998–1.000), followed by that associated with the LR model (0.768; 95% CI 0.729–0.808). When applied to the testing set, the ANN model produced the highest AUC value (0.761; 95% CI 0.634–0.888; Figure 1). The RF model also achieved an acceptable AUC value (0.735; 95% CI 0.616–0.854), while the AUC for the LR model was 0.668 (95% CI 0.480–0.857).

As shown in Figure 2, SHAP analysis on the ANN model showed that presence of a distal aneurysm and treatment modality were the most important features associated with periprocedural complications. These were also identified as important features for the RF model by feature importance analysis. Aneurysm size was one of the top features for all three ML models. Overall, we identified distal aneurysm, aneurysm size, and treatment modality as important features associated with endovascular treatment.

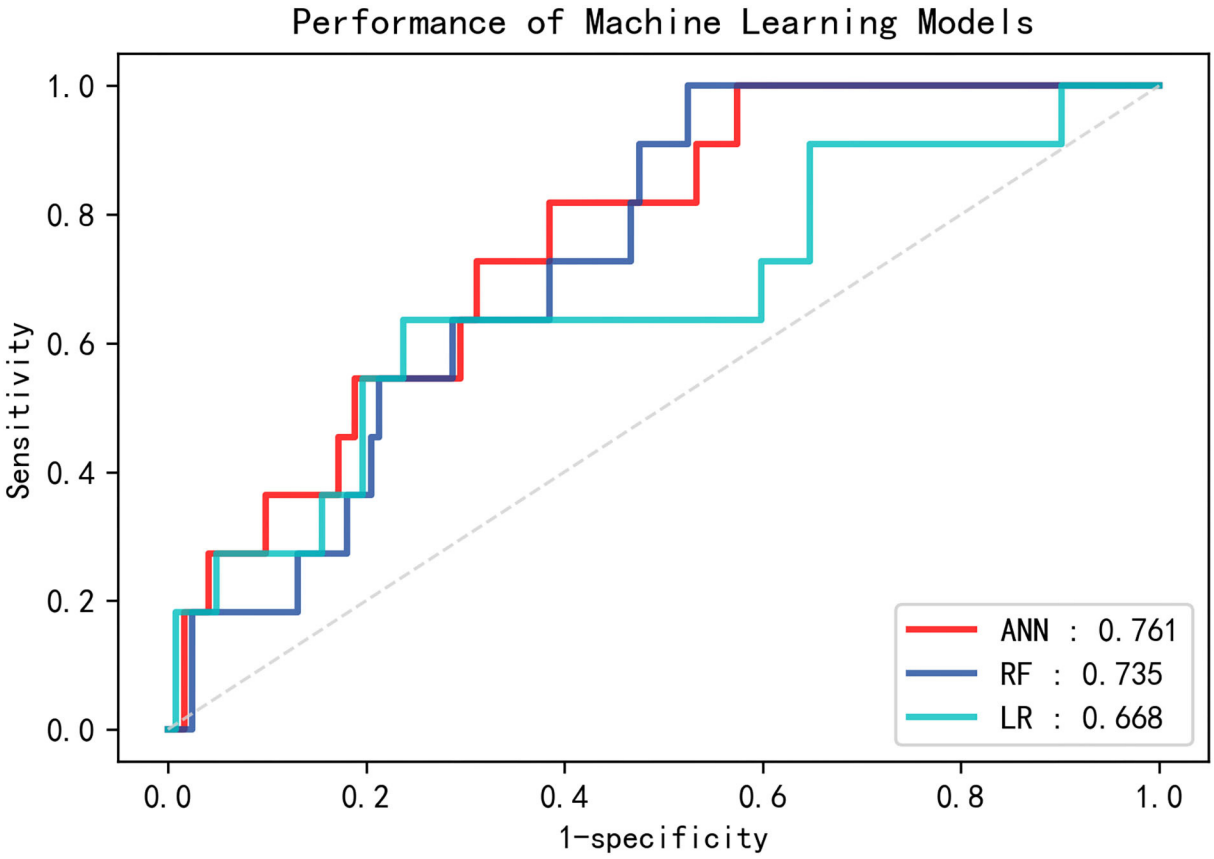


FIGURE 1
Receiver operating characteristic (ROC) curves for the three machine learning models (ANN, RF, and LR) on the testing set. ANN, artificial neural network; RF, random forest; LR, logistic regression.

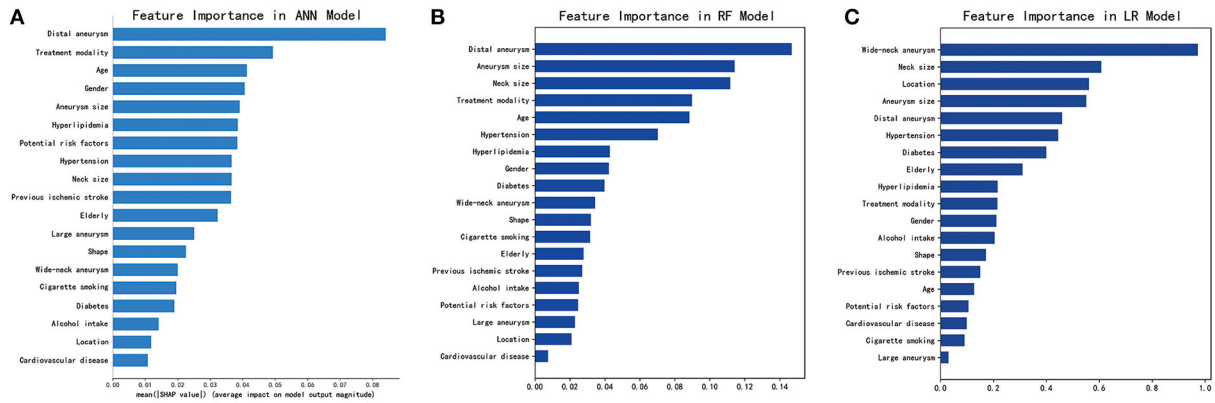


FIGURE 2
Identification of important features for the three machine learning models. **(A)** SHAP analysis for the ANN model. **(B)** Feature importance analysis for the RF model. **(C)** Feature importance analysis for the LR model. ANN, artificial neural network; RF, random forest; LR, logistic regression. For categorical variables, gender (male), potential risk factors (≥2), location (posterior circulation), and shape (irregular) were analyzed as potential risk factors. Receiver operating characteristic (ROC) curves for the three machine learning models (ANN, RF, and LR) on the testing set. ANN, artificial neural network; RF, random forest; LR, logistic regression.

TABLE 1 Results from univariate logistic regression analysis for all variables.

Characteristics	Control group ($N = 395$)	Complication group ($N = 48$)	*P*-value
Age (y)	55.59 ± 11.25	59.13 ± 12.28	0.043
Elderly	64 (16.2)	17 (35.4)	0.002
Gender (%)			0.276
Male	141 (35.7)	21 (43.8)	
Female	254 (64.3)	27 (56.3)	
Cigarette smoking (%)	41 (10.4)	6 (12.5)	0.653
Alcohol intake (%)	36 (9.1)	7 (14.6)	0.232
Hypertension (%)	195 (49.4)	34 (70.8)	0.006
Hyperlipidemia	108 (27.3)	13 (27.1)	0.970
Cardiovascular disease	17 (4.3)	2 (4.2)	0.965
Diabetes	35 (8.9)	2 (4.2)	0.279
Previous ischemic stroke	106 (26.8)	17 (35.4)	0.212
Potential risk factors (≥2)	165 (41.8)	27 (56.3)	0.058
Aneurysm size (mm)	6.71 ± 4.98	8.67 ± 5.54	0.014
Large aneurysm (%)	66 (16.7)	15 (31.3)	0.016
Neck size (mm)	5.21 ± 3.82	5.80 ± 2.61	0.301
Wide-neck aneurysm	349 (88.4)	47 (97.6)	0.075
Shape (%)			0.731
Regular	305 (77.2)	36 (75.0)	
Irregular	90 (22.8)	12 (25.0)	
Location (%)			0.906
Anterior circulation	364 (92.2)	44 (91.7)	
Posterior circulation	31 (7.8)	4 (8.3)	
Distal aneurysm	109 (27.6)	20 (41.7)	0.045
Treatment modality (%)			
Coiling	70 (17.7)	5 (10.4)	Ref
Stent-assisted coiling	240 (60.8)	30 (62.5)	0.265
Flow diverter	85 (21.5)	13 (27.1)	0.167

Discussion

Periprocedural complications associated with endovascular treatment for UIA represent a source of serious concern for practitioners. In the current study, we developed three ML models to predict these events and investigate risk factors associated with periprocedural complications. First and foremost, our results demonstrate that it is feasible to predict periprocedural complications associated with endovascular treatment using ML. Distal aneurysm, aneurysm size, and treatment modality may be key risk factors associated with endovascular treatment. Our findings may serve as a reference for physicians, and aid their decision-making process prior to UIA treatment.

ML is advantageous in exploring complex non-linear relationships across large datasets, and is a promising tool for clinical decision-making (17). Although many studies have reported successful ML prediction of risk for aneurysm rupture, there is little research on the application of ML to the prediction of periprocedural complications associated with endovascular treatment (18, 19). Ji et al. developed a scoring system for predicting the risk of neurological complications after endovascular treatment of UIAs, but their system was based on only three key factors (aneurysm size, aneurysm location, and cerebral ischemic comorbidity). Their approach may therefore be unsuitable for real-world applications (20). Staartjes et al. explored the feasibility of predicting neurological deficits after microsurgery for UIAs *via* application of ML techniques, and found that these methods support adequate prediction of early clinical endpoints after microsurgery for UIAs (21). However, their study did not include endovascularly treated UIA patients, and their models may therefore perform poorly when applied to such cases. In this study, we developed three ML models to predict perioperative complications associated with endovascular treatment for UIAs. Our results show that the ANN and RF models deliver satisfactory performance, indicating that ML is a valuable tool for prediction of perioperative complications after endovascular treatment for UIAs.

Distal aneurysm is an important predictor of periprocedural complications. In distal aneurysms, diameter of the parent artery and aneurysm size are often relatively small (22). In addition, the parent artery often presents several anatomical variants with numerous perforators or important small vessels that cannot be displayed on digital subtraction angiography. At the same time, sacrificing such vessels can result in neurological deficits (20), and distal location can increase arterial tortuosity. These factors pose serious challenges for successful endovascular treatment. Furthermore, they restrict the movement of endovascular devices, thus resulting in a higher rate of periprocedural complications. In this study, distal aneurysm was the most important feature for both ANN and RF models.

The modality of endovascular treatment has been demonstrated to be closely associated with periprocedural complications. Piotin et al. reported results from 1137 patients treated by coiling only or stent-assisted coiling (23). These authors found that stent-assisted coiling caused more permanent neurologic complications than coiling only (7.4 vs. 3.8%, $p = 0.64$) and a higher procedure-related mortality (4.6 vs. 1.2%, $p = 0.006$). Algra et al. reported that stents are associated with a higher complication risk than coiling (24). Naggara

et al. found that the use of a flow-diverter device doubled the risk of unfavorable outcomes compared with simple coil placement (25). In accordance with these previous studies, we found that coiling was the safest treatment modality. Compared with coiling only, both flow-diverter devices and stent-assisted coiling resulted in more periprocedural complications. We also found that treatment modality was one of the most important features for both ANN and RF models, although this result did not reach statistical significance after univariate logistic regression analysis. This apparent discrepancy may be due to the advantage of ML over conventional statistical methods in dealing with complex non-linear relationships across large datasets.

Larger aneurysm size has been reported to be associated with increased risk of periprocedural complications after endovascular treatment (26). Larger aneurysm size increases the complexity of endovascular procedures, and impedes good wall apposition for stent deployment (27). Furthermore, the embolization rate of intracranial aneurysm decreases with increasing aneurysm size, which means that larger aneurysms are more likely to carry residual flow within the coil mass (28–30). Our results confirm and extend these findings by demonstrating that aneurysm size is larger in the complication group compared with the control group. Furthermore, we found that aneurysm size was an important feature for all three ML models.

Limitations

Our study presents several limitations. Our dataset is relatively small and may involve patient selection bias. Therefore, our results may not generalize well to other patients and settings. Moreover, the synthetic data generated by the ADASYN procedure may not adequately represent less frequent cases. Future verification of our findings and validation of our models will require larger datasets from multiple centers.

Conclusion

Periprocedural complications after endovascular treatment for UIA can carry substantial consequences for patients. We show that these complications can be successfully predicted using ML models. These models represent promising tools for aiding decision-making prior to UIA treatment.

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 Ethics Committee of Zhujiang Hospital of Southern Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

ZT performed the manuscript writing and statistical analysis. WL processed the data. XF and KS acquired the data. CD conceived and designed the research. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the National Key Research Development Program (Grant numbers: 2016YFC1300804 and 2016YFC1300800), the National Natural Science Foundation of China (81974178, 81974177, and 82001300), the Science and Technology Project Foundation of Guangdong province (Grant number: 2016A020215098), and the Key Project of Clinical Research of Southern Medical University (Grant number: LC2016ZD024).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2022.1027557/full#supplementary-material

References

1. Vlak MH, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol.* (2011) 10:626–36. doi: 10.1016/S1474-4422(11)70109-0

2. Li MH, Chen SW, Li YD, Chen YC, Cheng YS, Hu DJ, et al. Prevalence of unruptured cerebral aneurysms in Chinese adults aged 35 to 75 years: a cross-sectional study. *Annal Int Med.* (2013) 159:514–21. doi: 10.7326/0003-4819-159-8-201310150-00004

3. Steiner T, Juvela S, Unterberg A, Jung C, Forsting M, Rinkel G. European Stroke Organization guidelines for the management of intracranial aneurysms and subarachnoid haemorrhage. *Cereb Dis.* (2013) 35:93–112. doi: 10.1159/000346087

4. Molyneux AJ, Kerr RS, Yu LM, Clarke M, Sneade M, Yarnold JA, et al. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion. *Lancet.* (2005) 366:809–17. doi: 10.1016/S0140-6736(05)67214-5

5. Malhotra A, Wu X, Forman HP, Matouk CC, Gandhi D, Sanelli P. Management of tiny unruptured intracranial aneurysms: a comparative effectiveness analysis. *JAMA Neurol.* (2018) 75:27–34. doi: 10.1001/jamaneurol.2017.3232

6. Park HK, Horowitz M, Jungreis C, Genevro J, Koebbe C, Levy E, et al. Periprocedural morbidity and mortality associated with endovascular treatment of intracranial aneurysms. *AJNR.* (2005) 26:506–14.

7. Xiong W, Chen T, Li J, Xiang L, Zhang C, Xiang L, et al. Interpretable machine learning model to predict rupture of small intracranial aneurysms and facilitate clinical decision. *Neurol Soc Italian Soc Clin Neurophysiol.* (2022). doi: 10.1007/s10072-022-06351-x. [Epub ahead of print].

8. Chen R, Mo X, Chen Z, Feng P, Li H. An Integrated Model combining machine learning and deep learning algorithms for classification of rupture status of IAs. *Front Neurol.* (2022) 13:868395. doi: 10.3389/fneur.2022.868395

9. Lin S, Zou Y, Hu J, Xiang L, Guo L, Lin X, et al. Development and assessment of machine learning models for predicting recurrence risk after endovascular treatment in patients with intracranial aneurysms. *Neurosurg Rev.* (2022) 45:1521–31. doi: 10.1007/s10143-021-01665-4

10. Liu Q, Jiang P, Jiang Y, Ge H, Li S, Jin H, et al. Prediction of aneurysm stability using a machine learning model based on pyradiomics-derived morphological features. *Stroke.* (2019) 50:2314–21. doi: 10.1161/STROKEAHA.119.025777

11. Zhu W, Li W, Tian Z, Zhang Y, Wang K, Zhang Y, et al. Stability assessment of intracranial aneurysms using machine learning based on clinical and morphological features. *Trans Stroke Res.* (2020) 11:1287–95. doi: 10.1007/s12975-020-00811-2

12. Paliwal N, Jaiswal P, Tutino VM, Shallwani H, Davies JM, Siddiqui AH, et al. Outcome prediction of intracranial aneurysm treatment by flow diverters using machine learning. *Neurosurg Focus.* (2018) 45:E7. doi: 10.3171/2018.8.FOCUS18332

13. Guédon A, Thépenier C, Shotar E, Gabrieli J, Mathon B, Premat K, et al. Predictive score for complete occlusion of intracranial aneurysms treated by flow-diverter stents using machine learning. *J Neurointerv Surg.* (2020) 13:341–6. doi: 10.1136/neurintsurg-2020-016748

14. Kallmes DF, Hanel R, Lopes D, Boccardi E, Bonafe A, Cekirge S, et al. International retrospective study of the pipeline embolization device: a multicenter aneurysm treatment study. *AJNR.* (2015) 36:108–15. doi: 10.3174/ajnr.A4111

15. Liu J, Chen Y, Lan L, Lin B, Chen W, Wang M, et al. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. *Eur Radiol.* (2018) 28:3268–75. doi: 10.1007/s00330-017-5300-3

16. Ou C, Liu J, Qian Y, Chong W, Zhang X, Liu W, et al. Rupture risk assessment for cerebral aneurysm using interpretable machine learning on multidimensional data. *Front Neurol.* (2020) 11:570181. doi: 10.3389/fneur.2020.570181

17. Marasini A, Shrestha A, Phuyal S, Zaidat OO, Kalia JS. Role of artificial intelligence in unruptured intracranial aneurysm: an overview. *Front Neurol.* (2022) 13:784326. doi: 10.3389/fneur.2022.784326

18. Chen G, Lu M, Shi Z, Xia S, Ren Y, Liu Z, et al. Development and validation of machine learning prediction model based on computed tomography angiography-derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. *Eur Radiol.* (2020) 30:5170–82. doi: 10.1007/s00330-020-06886-7

19. Shi Z, Chen GZ, Mao L, Li XL, Zhou CS, Xia S, et al. Machine learning-based prediction of small intracranial aneurysm rupture status using CTA-derived hemodynamics: a multicenter study. *AJNR.* (2021) 42:648–54. doi: 10.3174/ajnr.A7034

20. Ji W, Liu A, Lv X, Kang H, Sun L, Li Y, et al. Risk Score for neurological complications after endovascular treatment of unruptured intracranial aneurysms. *Stroke.* (2016) 47:971–8. doi: 10.1161/STROKEAHA.115.012097

21. Staartjes VE, Sebök M, Blum PG, Serra C, Germans MR, Krayenbühl N, et al. Development of machine learning-based preoperative predictive analytics for unruptured intracranial aneurysm surgery: a pilot study. *Acta Neurochirurgica.* (2020) 162:2759–65. doi: 10.1007/s00701-020-04355-0

22. Park YK, Yi HJ, Choi KS, Lee YJ, Chun HJ. Intraprocedural rupture during endovascular treatment of intracranial aneurysm: clinical results and literature review. *World Neurosurg.* (2018) 114:e605–15. doi: 10.1016/j.wneu.2018.03.040

23. Piotin M, Blanc R, Spelle L, Mounayer C, Piantino R, Schmidt PJ, et al. Stent-assisted coiling of intracranial aneurysms: clinical and angiographic results in 216 consecutive aneurysms. *Stroke.* (2010) 41:110–5. doi: 10.1161/STROKEAHA.109.558114

24. Algra AM, Lindgren A, Vergouwen MDI, Greving JP, van der Schaaf IC, van Doormaal TPC, et al. Procedural clinical complications, case-fatality risks, and risk factors in endovascular and neurosurgical treatment of unruptured intracranial aneurysms: a systematic review and meta-analysis. *JAMA Neurol.* (2019) 76:282–93. doi: 10.1001/jamaneurol.2018.4165

25. Naggara ON, Lecler A, Oppenheim C, Meder JF, Raymond J. Endovascular treatment of intracranial unruptured aneurysms: a systematic review of the literature on safety with emphasis on subgroup analyses. *Radiology.* (2012) 263:828–35. doi: 10.1148/radiol.12112114

26. Orrù E, Roccatagliata L, Cester G, Causin F, Castellan L. Complications of endovascular treatment of cerebral aneurysms. *Eur J Radiol.* (2013) 82:1653–8. doi: 10.1016/j.ejrad.2012.12.011

27. Brinjikji W, Lanzino G, Cloft HJ, Siddiqui AH, Boccardi E, Cekirge S, et al. Risk factors for ischemic complications following pipeline embolization device treatment of intracranial aneurysms: results from the intreped study. *AJNR.* (2016) 37:1673–8. doi: 10.3174/ajnr.A4807

28. Ihn YK, Shin SH, Baik SK, Choi IS. Complications of endovascular treatment for intracranial aneurysms: Management and prevention. *Int Neuroradiol J Perither Neuroradiol Surg Proc Neurosci.* (2018) 24:237–45. doi: 10.1177/1591019918758493

29. Pierot L, Cognard C, Anxionnat R, Ricolfi F. Ruptured intracranial aneurysms: factors affecting the rate and outcome of endovascular treatment complications in a series of 782 patients (CLARITY study). *Radiology.* (2010) 256:916–23. doi: 10.1148/radiol.10092209

30. Tamatani S, Ito Y, Koike T, Abe H, Kumagai T, Takeuchi S, et al. Evaluation of the stability of intracranial aneurysms occluded with guglielmi detachable coils. *Int Neuroradiol J Perither Neuroradiol Surg Proc Neurosci.* (2001) 7:143–8. doi: 10.1177/15910199010070S121

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 26 April 2023
DOI 10.3389/fneur.2023.1095266

OPEN ACCESS

EDITED BY
Ferdinand Hui,
Johns Hopkins Medicine, United States

REVIEWED BY
Qinghai Huang,
Changhai Hospital, China
Philipp Gruber,
Aarau Cantonal Hospital, Switzerland

*CORRESPONDENCE
Ming-Xin Zhu
✉ mingxinzhu2017@163.com
Ting Lei
✉ tinglei_tj@126.com
Jian-Ping Xiang
✉ jianpingxiang1979@163.com

RECEIVED 11 November 2022
ACCEPTED 23 March 2023
PUBLISHED 26 April 2023

CITATION
Wu Z-B, Zeng Y, Zhang H-Q, Shu K, Li G-H, Xiang J-P, Lei T and Zhu M-X (2023) Virtual simulation with AneuShape™ software for microcatheter shaping in intracranial aneurysm coiling: a validation study.
Front. Neurol. 14:1095266.
doi: 10.3389/fneur.2023.1095266

Virtual simulation with AneuShape™ software for microcatheter shaping in intracranial aneurysm coiling: a validation study

Zeng-Bao Wu[1], Ying Zeng[1], Hua-Qiu Zhang[1], Kai Shu[1], Gao-Hui Li[2], Jian-Ping Xiang[2]*, Ting Lei[1]* and Ming-Xin Zhu[1]*

[1]Department of Neurosurgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [2]ArteryFlow Technology Co., Ltd., Hangzhou, Zhejiang, China

Background: The shaping of an accurate and stable microcatheter plays a vital role in the successful embolization of intracranial aneurysms. Our study aimed to investigate the application and the role of AneuShape™ software in microcatheter shaping for intracranial aneurysm embolization.

Methods: From January 2021 to June 2022, 105 patients with single unruptured intracranial aneurysms were retrospectively analyzed with or without AneuShape™ software to assist in microcatheter shaping. The rates of microcatheter accessibility, accurate positioning, and stability for shaping were analyzed. During the operation, fluoroscopy duration, radiation dose, immediate postoperative angiography, and procedure-related complications were evaluated.

Results: Compared to the manual group, aneurysm-coiling procedures involving the AneuShape™ software exhibited superior results. The use of the software resulted in a lower rate of reshaping microcatheters (21.82 vs. 44.00%, $p = 0.015$) and higher rates of accessibility (81.82 vs. 58.00%, $p = 0.008$), better positioning (85.45 vs. 64.00%, $p = 0.011$), and higher stability (83.64 vs. 62.00%, $p = 0.012$). The software group also required more coils for both small (<7 mm) and large (≥7 mm) aneurysms compared to the manual group (3.50 ± 0.19 vs. 2.78 ± 0.11, $p = 0.008$ and 8.22 ± 0.36 vs. 6.00 ± 1.00, $p = 0.081$, respectively). In addition, the software group achieved better complete or approximately complete aneurysm obliteration (87.27 vs. 66.00%, $p = 0.010$) and had a lower procedure-related complication rate (3.60 vs. 12.00%, $p = 0.107$). Without this software, the operation had a longer intervention duration (34.31 ± 6.51 vs. 23.87 ± 6.98 min, $p < 0.001$) and a higher radiation dose (750.50 ± 177.81 vs. 563.53 ± 195.46 mGy, $p < 0.001$).

Conclusions: Software-based microcatheter shaping techniques can assist in the precise shaping of microcatheters, reduce operating time and radiation dose, improve embolization density, and facilitate more stable and efficient intracranial aneurysm embolization.

KEYWORDS

virtual simulation, microcatheter shaping, cerebral aneurysm, coil embolization, treatment outcome

1. Introduction

The precise insertion and stabilization of a microcatheter within an aneurysmal sac are crucial for a successful interventional embolization procedure (1–3). Therefore, it is essential to properly shape the microcatheter to achieve optimal navigation and stability (4). Inappropriate microcatheter shaping can cause the tip to rebound from the aneurysmal sac prematurely, obstructing further packing of the coil. Microcatheter shaping is not an easy process for neurosurgeons, even though it is a routine technique for the interventional embolization of intracranial aneurysms. Currently, microcatheter shaping mainly relies on the individual experience and estimation of operators and the visualization of the path of the microcatheter in the aneurysmal cavity and the parent artery according to digital subtraction angiography (DSA) (5). However, for aneurysms located in tricky areas or with intricate anatomical morphology, even sophisticated neurosurgeons may need to reshape the microcatheter or make multiple attempts to achieve the right shape to access the aneurysm sac (6). Some new methods of microcatheter formation have been reported in the literature; however, there are different limitations (6–8). The AneuShape™ software is a real-time planning tool that provides the microcatheter shaping template before the operation and simulates the path of the shaped microcatheter in the aneurysm sac and the parent artery with high accuracy, assisting neuro-interventionalists with accurate microcatheter shaping. In this research, we investigated the safety, accuracy, and effectiveness of using AneuShape™ software for microcatheter shaping during the endovascular embolization of intracranial aneurysms and also whether they were superior to traditional manual shaping methods.

2. Methods

2.1. Study participants

This was a retrospective single-institution series that was authorized by the institutional medical ethics committee. From January 2021 to June 2022, 105 consecutive patients with 105 single unruptured intracranial aneurysms in the anterior cerebrovascular circulation underwent endovascular therapy in our center. Neuro-interventionalists with 3–5 years of interventional surgery experience were in charge of the entire procedure. Software-based microcatheter shaping was applied in 55 cases, while conventional manual microcatheter shaping was applied in 50 cases. Collected data included demographics, aneurysm characteristics, treatment duration, radiation dose, and complications. The rates of microcatheter reshaping and accessibility, accurate positioning, and stability for shaping were also analyzed. All interventional therapies utilized coiling embolization. Meanwhile, the size of the aneurysm and personal intraoperative conditions determined whether stent-assisted coiling embolization was necessary.

2.2. Inclusion criteria

The inclusion criteria were as follows: (1) the presence of unruptured aneurysms specifically located in the anterior cerebrovascular circulation, (2) the absence of any surgical-related contraindications, and (3) the patient's voluntary consent to use AneuShape™ software.

2.3. Exclusion criteria

The exclusion criteria for this study comprised the following parameters: (1) giant aneurysms, ruptured aneurysms, and aneurysms located in the posterior cerebrovascular circulation and (2) the use of double or multiple microcatheter coiling technology.

2.4. Endovascular procedure

Before the endovascular procedure, the related clinical data were obtained from the patient's medical records, and all patients provided written informed consent. General anesthesia was used in all patients. The femoral sheath was the site of catheter placement. To maintain an active clotting time of ≥250 s, the patients were given one dose of standard heparin (70–100 IU/kg), which was subsequently administered in additional hourly doses (1,000 IU). A 6 French Navien (Stryker, USA) or guiding catheter (Envoy; Johnson & Johnson Health Care Systems Inc., USA) was advanced to the petrous horizontal segment of the internal carotid artery. Afterward, the shaped microcatheter was inserted into the aneurysm sac through a guiding catheter or Navien with microguidewire guidance. Then, the aneurysm sac was densely embolized with coils. For cases requiring stent assistance, the endovascular procedure utilized the "coil-through" or "stent semi-jailing" technique.

2.5. Antiplatelet therapy protocol

Before the unruptured aneurysm was treated with a stent, dual antiplatelet treatment with aspirin (100 mg/d) and clopidogrel (75 mg/d) was administered for 3–5 days. In addition, tirofiban (Gland Pharma Limited, China) at 0.1 μg/kg of body mass/min was administered by continuous intravenous infusion pump for 12 h. After the embolization procedure, patients who received stent assistance were advised to maintain dual antiplatelet therapy daily for at least 3 months and 100 mg of aspirin for a minimum of 6 months.

2.6. Virtual simulation with AneuShape™ software for microcatheter shaping

AneuShape™ (ArteryFlow Technology, Hangzhou, China) is a real-time planning tool to assist neuro-interventionalists in microcatheter shaping. The workflow of AneuShape™ begins with reading the preoperative 3D rotational angiography (3DRA) images (Figure 1A). Based on level-set segmentation, the intracranial vasculature is then reconstructed, and a region of interest (ROI), for example, the left paraclinoid aneurysms of the internal carotid artery and parent artery, can be manually

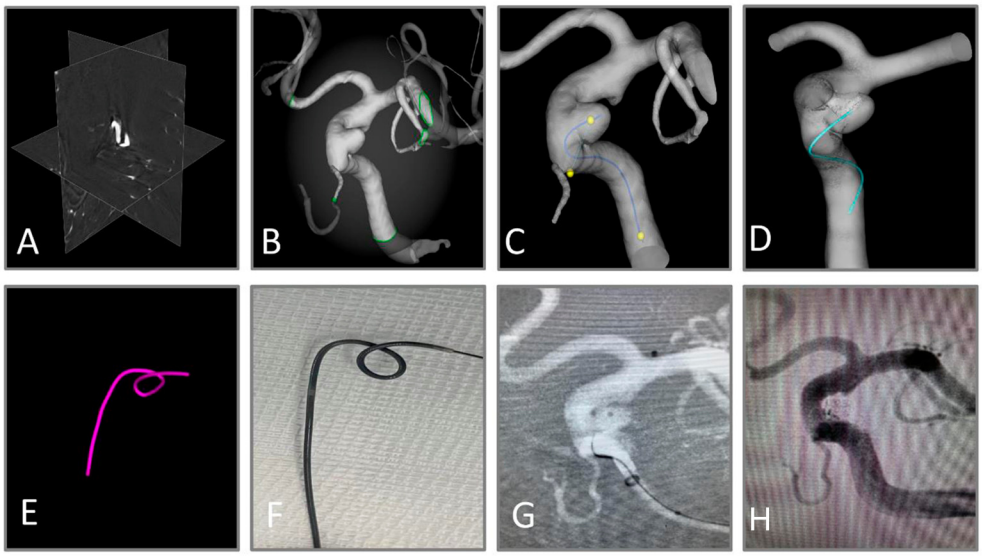


FIGURE 1
Workflow of virtual shaping of microcatheter with AneuShape™. **(A)** Reading of 3D RA images. **(B)** Segmentation and reconstruction of intracranial vessels. **(C)** Generation of the centerline, and the picking of three key points. **(D, E)** Visualization of anticipated microcatheter after removal of the mandrel (cyan) and that with mandrel inserted after shaping (purple). **(F)** The microcatheter was shaped manually using a template generated by the software. **(G)** The aneurysm sac was successfully catheterized through the shaped microcatheter. **(H)** Satisfactory aneurysm embolization.

extracted with a cropping sphere (Figure 1B). To determine the direction of the virtual catheter, a centerline starting from the proximal inlet to the aneurysm sac should be generated (Figure 1C), which is accomplished by clicking at the proximal inlet and the aneurysm dome. Then, doctors need to pick two points on the centerline representing the proximal location and the tip of the anticipated microcatheter pathway, as well as a third point that corresponds to the first contacting point between the virtual microcatheter and the vessel wall along the distal-to-proximal direction from the tip (Figure 1C). The virtual microcatheter will then appear in cyan in only a few seconds (Figure 1D). If the surgeon defines the shaping factor (which describes the blunting behavior of the microcatheter after shaping, e.g., a 90°-shaping angle with an inserted mandrel turning to 45° after the removal of the mandrel if this factor equals 2.0) and a shaping length, a purple tube will appear in real-time, showing the shaped path of the microcatheter after inserting the mandrel (Figure 1E).

The principle of virtual microcatheter simulation includes a collision detection algorithm and a direction correction algorithm. At first, the virtual microcatheter "grows" from the tip location (usually at the centroid of the sac) and is picked up on the centerline toward the first contacting point. The "growing" process indicates that trial points are generated one by one along the direction defined by the previous two points. Meanwhile, collision detection is performed for each trial point, which is based on the widely used ray-casting algorithm. If the current trial point is outside the vessel wall, then its nearest neighbor on the centerline is located. The connecting line between the current trial point and its nearest neighbor intersects with the vessel wall, and the former is then moved to this intersecting location (and the "growing" direction is corrected correspondingly) to keep it inside the vessel wall. Then, a new trial point is generated along the new direction, and the aforementioned process repeats until the trial point reaches the proximal location picked on the centerline. Finally, all the points are connected sequentially. The virtual microcatheter is a set of adjacent line segments and is visualized as a winding tube in the vessel lumen.

To convert the microcatheter pathway in the vessel to its shaped pathway, we calculated the rotation matrix between each two adjacent line segments. The rotation matrix maps a vector rigidly to a new vector by taking a rotating axis and a rotating angle as inputs. The rotating axis between two line segments is the cross-product of their direction vectors. The rotating angle is obtained by multiplying the angle between the direction vectors of the two adjacent line segments with a coefficient (which equals the shaping factor minus one). Based on the rotation matrix, each line segment is rotated one by one with respect to its proximal neighbor until the desired shaping length is reached.

2.7. Manual microcatheter shaping

By utilizing a template generated by software, the microcatheter was manually molded and steamed for 60 s to complete the shaping (Figure 1F). The crawling path of the virtual microcatheter and the shaped path together facilitated the neuro-interventionalists in optimizing the surgical plan. The aneurysm sac was catheterized in one attempt by the shaped microcatheter (Figure 1G). The aneurysm was embolized with coils until satisfactory saccular obliteration was accomplished (Figure 1H).

2.8. Microcatheter evaluation and data collection

Two experienced neuro-interventionists determined whether it was "good" or "poor" in assessing the accessibility, the in-position condition, and the stability of each shaped microcatheter during procedures (9). Whether the microcatheters needed reshaping was also recorded. If the aneurysm sac was catheterized in 5 m, the accessibility of the microcatheter was defined as "good." The position was defined as "poor" if the microcatheter tip was adherent within the intracranial aneurysm sac. However, if the microcatheter prematurely retreated out of the aneurysm sac and affected the further packing of the coil, stability was defined as "poor." The fluoroscopy duration was measured in minutes and was defined as the time required from the beginning, when the microcatheter was out of the guiding catheter, to successful entry into the aneurysm sac and the time taken to release all coils. However, the time required to deploy the guiding catheter and stent was excluded. If there were multiple attempts at microcatheter delivery, the time for all attempts was recorded. The radiation dose was defined as the dose generated over the duration of treatment in terms of air kerma in mGy. We defined paraclinoid artery and posterior communicating artery aneurysms as proximal and middle cerebral artery bifurcations and anterior communicating artery aneurysms as distal aneurysms. The difference in efficiency between the distal and proximal aneurysms using software-assisted microcatheter shaping was also compared.

The aneurysm embolization rate and complications were also evaluated and noted by two independent neuro-interventionists. We applied the modified Roy-Raymond classification (MRRC) to assess the immediate angiographic results, and the classification criteria include (9): Class I (complete obliteration), Class II (neck residual), and Class III (non-complete occlusion). Our research designated Class I and Class II as successful intracranial aneurysm embolization.

2.9. Statistical analysis

Continuous data were expressed as the mean ± standard deviation (SD) and compared using the *t*-test, which conformed to a normal distribution, while the Mann–Whitney *U*-test was used for data that no longer fit the normal distribution. Categorical data were expressed as numbers (percentages) and compared using the chi-squared test. A *p*-value of <0.05 was defined as statistically significant. The SPSS 22.0 software (IBM, USA) was used for statistical analysis.

3. Results

3.1. Patients

A total of 105 cases involved 50 patients who were not treated with software-assisted technology and 55 patients treated with software-based technology. Table 1 shows the baseline characteristics of cases treated with and without software-based technology. The mean age of patients was 56.71 ± 8.95 years. There were 71 (67.62%) sidewall aneurysms and 34 (32.38%) bifurcation aneurysms, which were all located in the anterior circulation. The mean maximum diameter, neck width, and dome/neck ratio of all aneurysms were 4.40 ± 2.30, 5.25 ± 1.99, and 0.79 ± 0.36 mm, respectively. Moreover, there were no statistical differences in morphological parameters between the two groups ($p > 0.05$; Table 1).

TABLE 1 Baseline clinical data and aneurysm characteristics in the software and manual groups.

Factor	With software $n = 55$	Without software $n = 50$	*p*-value
Age, years mean ± SD	57.16 ± 9.30	56.22 ± 8.63	0.592
Male sex, *n* (%)	15 (27.27)	16 (30.0)	0.596
Dmax, mm mean ± SD	4.58 ± 2.64	4.21 ± 1.88	0.89
Neck, mm mean ± SD	5.23 ± 2.05	5.27 ± 1.95	0.90
Size (%)			0.194
Small (<7 mm)	46 (83.60)	46 (92.00)	
Large (≥7 mm)	9 (16.40)	4 (8.00)	
AR	0.80 ± 0.35	0.78 ± 0.41	0.579
SR	1.57 ± 1.03	1.38 ± 0.88	0.49
Sidewall/bifurcation aneurysm (%)			0.937
Sidewall, *n* (%)	37 (67.27)	34 (68.0)	
Bifurcation, *n* (%)	18 (32.73)	16 (32.0)	
Location (%)			0.983
Paraclinoid aneurysms, *n* (%)	20 (36.30)	19 (38.00)	
PcomA aneurysms, *n* (%)	17 (30.90)	15 (30.00)	
AcomA aneurysms, *n* (%)	9 (16.40)	7 (14.00)	
MCA-Bifurcation aneurysms, *n* (%)	9 (16.40)	9 (18.00)	
Treatment therapy (%)			0.678
Coiling, *n* (%)	14 (25.50)	11 (22.00)	
Stent-assisted coiling, *n* (%)	41 (74.50)	39 (78.00)	

SD, standard deviation; AR, aspect ratio; SR, size ratio; PcomA, posterior communicating artery aneurysm; AcomA, anterior communicating artery aneurysm; MCA, middle cerebral artery.

3.2. Initial outcomes

In the software group, three aneurysms required an adjunctive procedure in which the microcatheter was discarded and a new one was reshaped for replacing the old one. However, in the group

TABLE 2 Number of coils used.

	Small aneurysms <7 mm			Large aneurysms ≥7 mm		
	With software $n = 46$	Without software $n = 46$	*p*-value	With software $n = 9$	Without software $n = 4$	*p*-value
Dmax, mm mean ± SD	3.59 ± 0.20	3.86 ± 0.21	0.419	9.63 ± 0.55	8.22 ± 0.94	0.178
Coils (*n*)	3.50 ± 0.19	2.78 ± 0.11	0.008	8.22 ± 0.36	6.00 ± 1.00	0.081

without software, 12 aneurysms required a manual procedure in which the microcatheter was discarded and a new one was shaped.

For both small (<7 mm) and large (≥7 mm) aneurysms, the software group used more coils than the manual group (3.50 ± 0.19 vs. 2.78 ± 0.11, $p = 0.008$ and 8.22 ± 0.36 vs. 6.00 ± 1.00, $p = 0.081$, respectively; Table 2).

In comparison with the manual group, the software group gained a shorter fluoroscopy duration (23.87 ± 6.98 vs. 34.31 ± 6.51 min, $p < 0.001$) and a lower radiation dose (563.53 ± 195.46 vs. 750.50 ± 177.81 mGy, $p < 0.001$) in terms of the univariate analysis (Table 3). In the coil-only group, the fluoroscopy duration and radiation dose were 16.88 ± 0.75 vs. 28.50 ± 1.16 min, $p < 0.001$, and 381.00 ± 18.83 vs. 597.27 ± 33.32 mGy, $p < 0.001$, respectively (Table 4). Software-based microcatheter shaping techniques also reduced operating time and radiation dose in the stent-assisted coiling group, which were 26.25 ± 0.99 vs. 35.95 ± 0.99 min, $p < 0.001$ and 625.85.00 ± 28.93 vs. 793.72 ± 27.19 mGy, $p < 0.001$, respectively (Table 4). Meanwhile, we found that there was a linear correlation between fluoroscopy duration and radiation dose (Spearman's correlation coefficient = 0.944, $p < 0.001$).

In addition, the software group gained a lower reshaping rate for microcatheters (21.82 vs. 44.00%, $p = 0.015$) and a higher rate of accessibility (81.82 vs. 58.00%, $p = 0.008$) and achieved a better positioning (85.45 vs. 64.00%, $p = 0.011$) and higher stability (83.64 vs. 62.00%, $p = 0.012$; Table 3). Compared with the proximal aneurysms, the efficiency of using software-assisted microcatheter shaping was not reduced, and the difference between the two groups was not statistically significant (Table 5).

Post-embolism angiograms showed that the software group gained a higher rate of complete occlusion (Raymond-Roy Grade Scale I [RRGS I]: 74.54 vs. 48.00%, $p = 0.005$), a lower rate of neck remnant (RRGS II: 12.73 vs. 18.00%, $p = 0.453$), and non-complete occlusion (RRGS III: 12.73 vs. 34.00%, $p = 0.010$; Table 3), as well as a greater degree of complete or approximately complete aneurysm occlusion (87.27 vs. 66.00%, $p = 0.010$).

TABLE 3 Outcomes and assessment of microcatheter shaping.

	With software $n = 55$	Without software $n = 50$	*p*-value
Outcomes			
Need for reshaping, *n* (%)	12 (21.82)	22 (44.00)	0.015
Duration of the intervention, mean ± SD (min)	23.87 ± 6.98	34.31 ± 6.51	<0.001
Radiation dose, mean ± SD (mGy)	563.53 ± 195.46	750.50 ± 177.81	<0.001
Postoperative angiography			
Raymond grade 1, *n* (%)	41 (74.54)	24 (48.00)	0.005
Raymond grade 2, *n* (%)	7 (12.73)	9 (18.00)	0.453
Raymond grade 3, *n* (%)	7 (12.73)	17 (34.00)	0.010
Complication, *n* (%)	2 (3.60)	6 (12.00)	0.107
Microcatheter shaping			
Accessibility, good, *n* (%)	45 (81.82)	29 (58.00)	0.008
Positioning, good, *n* (%)	47 (85.45)	32 (64.00)	0.011
Stability, good, *n* (%)	46 (83.64)	31 (62.00)	0.012

3.3. Perioperative complications

In the manual group, six patients (12.00%) experienced various complications: three minor strokes with mild motor symptoms degenerating, one small frontal hematoma with a mild headache, one asymptomatic internal carotid artery dissection, and one regional obliteration of the right middle cerebral artery with partly resolved hemiparesis. In the group with software, two patients (3.60%) experienced complications: one patient had a mild ischemic stroke with transient muscle weakness of the lower limbs, and another had a stroke with mild right hemiparesis and dysarthria. All complications except one (the middle cerebral artery partial occlusion in the manual group that caused permanent weakness of the upper limbs) completely reverted without sequelae.

4. Discussion

The optimal microcatheter shape has a significant impact on retaining the accuracy and stability of the microcatheter during embolization, enabling safe, and effective treatment of intracranial aneurysms (1, 2). In our series, the manual and software groups were compared. Microcatheter shaping based on the AneuShape™ software had reliable stability and accuracy in endovascular treatment and played a significant role in the successful embolization of intracranial aneurysms. Besides, the use of the real-time preoperative virtual shaping technique was associated with a significant reduction in the intervention duration

TABLE 4 Software-based microcatheter shaping techniques reduce operating time and radiation dose both in coiling-only group and stent-assisted coiling group.

	Coiling			Stent-assisted coiling		
	With software $n = 14$	Without software $n = 11$	*p*-value	With software $n = 41$	Without software $n = 39$	*p*-value
Dmax, mm mean ± SD	2.20 ± 0.12	2.25 ± 0.14	0.766	5.39 ± 0.40	4.76 ± 0.28	0.45
Duration of the intervention, mean ± SD (min)	16.88 ± 0.75	28.50 ± 1.16	<0.001	26.25 ± 0.99	35.95 ± 0.99	<0.001
Radiation dose, mean ± SD (mGy)	381.00 ± 18.83	597.27 ± 33.32	<0.001	625.85.00 ± 28.93	793.72 ± 27.19	<0.001

TABLE 5 Comparison of microcatheter shaping efficiency of proximal and distal aneurysms by software-assisted.

With software	Proximal aneurysms $n = 37$	Distal aneurysms $n = 18$	*p*-value
Duration of the intervention, mean ± SD (min)	23.18 ± 6.39	25.28 ± 8.08	0.300
Radiation dose, mean ± SD (mGy)	552.70 ± 190.60	585.78 ± 209.00	0.531
Accessibility, good, *n* (%)	31 (83.78)	14 (77.78)	0.866
Positioning, good, *n* (%)	32 (86.49)	15 (83.33)	1
Stability, good, *n* (%)	31 (83.78)	15 (83.33)	1

and the radiation dose in both the coiling-only group and the stent-assisted coiling group.

Currently, most neuro-interventionalists use 3D-DSA images to evaluate the anatomy of aneurysms and their correlation to the parent artery before manually shaping the microcatheter. While microcatheter shaping is an essential technique for interventional treatment, it is sometimes difficult to obtain a satisfactory shape. One reason is that 3D imaging technology lacks depth information; hence, neuro-interventionalists cannot accurately perceive the 3D cerebrovascular morphology and the spatial position of microcatheters in that parent artery (4, 6). Second, it is hard for neuro-interventionalists to accurately identify the real route of the microcatheter into the parent artery, making them rely solely on their imagination. According to the characteristics of the microcatheter, neuro-interventionalists can adjust it into different shapes, but exaggerating the mandrel shape during the operation still depends on personal experience (6, 10). However, even experienced surgeons cannot guarantee that the shaping of the microcatheter is always appropriate, and sometimes reshaping is needed.

Furthermore, a recent multicenter study has shown that interventionists often use animal models for simulation training, and there is a lack of standardized training in neurological interventions (11). Hence, there is a need to develop a new artificial method that meets the training requirements of alternative animal models, which can increase the experience of neuro-interventionists faster. Software-based technology provides a promising method for interventional surgery simulation training and direct utilization of software-based technology during surgery.

Unlike conventional manual shaping processes that rely primarily on personal experience, the software can input DICOM data directly from 3D-DSA images during operation to generate microcatheter shaping templates. This makes it more accessible and convenient for doctors to mold the microcatheter at once and avoids repeated manual shaping, especially for less experienced neurosurgeons. In addition, software-based technology also offers a standardized manual training system for beginners, allowing neurosurgeons to quickly obtain intuitive experience and skills (12). In related research (12, 13), some researchers have proposed that the use of software-based technology to help surgeons shape the mandrel in the process of embolization of intracranial aneurysms achieves good clinical results. Liu et al. (12) reported that during endovascular surgery, the software-based simulation template entered the aneurysm sacs accurately and without any complications. Software-based microcatheter shaping was stable and accurate. However, these studies lacked a control group to assess the technical and clinical outcomes of software-based technology vs. traditional manual microcatheter shaping.

In our study, software-based microcatheter shaping was more accurate and stable compared with the manual shaping process. Therefore, we observed a higher rate of accessibility, achieving a better position and higher stability. In addition, the use of software-based technology reduced the need for reshaping the microcatheter or deploying a second one because of fatigue caused by repeated use. These findings were seldom quantified and reported in past research. These factors not only resulted in longer operating times and unproductive costs but also posed potential risks for procedural complications.

The reduction in duration with the software-based technology is likely related to the precise shaping of the microcatheter. As for complicated structures, neuro-interventionalists should perhaps remodel the microcatheter and make multiple attempts during the procedure to adapt the structural characteristics of the aneurysm sac and the parent artery. These procedures can prolong the procedural duration and raise the risk of aneurysm rupture and ischemic events (14–16). Software-based technology can reduce the need for repeated and inefficient maneuvers, making the procedure

smoother and more efficient. Surgeons can obtain the template for microcatheter shaping with pivotal data, such as the angle and length of the microcatheter tip, in ∼10 min. In our research, the operating time of the software-based group was reduced by about 30.43% compared to the manual group. Despite the reduced duration of the intervention, we found an increase in the number of coils used with the help of software-assisted shaping, implying denser aneurysm embolisms and lower recurrence rates. Our research results also showed that the software-based group achieved more complete or approximately complete aneurysm occlusions.

Longer operative times may be related to higher ischemic stroke rates during endovascular therapy (14, 15) or diagnostic angiography (16). It has been reported that the intervention treatment over 100–120 min was significantly associated with the risk of ischemic stroke events (14, 15). Similar results were shown in our study. Compared with the manual group, the complication rate in the software-based technology group was lower. Although the difference in the complication rate between those two groups was not significant, it appears plausible that a shorter operation time might reduce the risk of silent or symptomatic ischemic events.

Another focus of our research was the reduction in radiation dose, which is linearly related to the duration of fluoroscopy. Appropriate software-based microcatheter shaping reduces procedure time and, thus, facilitates a lower radiation dose. This is beneficial in reducing radiation-related risks for both patients and neuro-interventionalists.

This study has several limitations. First, this study used a retrospective design, which may have cause an introduced bias. Second, the shaping of the microcatheter by the surgeon according to the software-based template could still be subject to personal biases. Third, the number of cases was not large enough, and the data represented the limited experience of our institution. Fourth, due to the absence of randomization between the groups, the uniform distribution of aneurysms could not be ensured. Five, the impact of software-based techniques on clinical outcomes could not be better compared due to the lack of long-term follow-up. Six, software-based technology can improve the success rate of microcatheter shaping for low- and mid-level neuro-interventionists with 3–5 years of experience, but it is not clear whether it will assist the success rate of senior doctors (more than 5 years of experience). Therefore, further research is needed to investigate the use of the software by senior neuro-interventionists. Finally, multicenter randomized controlled trials are required to verify the effectiveness of this technology in the future.

5. Conclusions

In our experience, software-based technology is a beneficial tool that can assist neuro-interventionists in accurately shaping microcatheters. The use of this software improved embolization density, increased the rate of successful aneurysmal occlusion, and reduced the need for rectification procedures, fluoroscopy time, and radiation dose.

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 Medical Ethics Committee of Tongji Hospital. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

Z-BW, TL, and M-XZ designed and conducted the experiments and statistical analysis and drafted the manuscript. YZ participated in the operation and the collection of clinical data. G-HL and J-PX were responsible for software processing and data curation. H-QZ and KS conducted intraoperative and postoperative evaluations. All authors contributed to the article and approved the submitted version.

Conflict of interest

G-HL and J-PX were employed by ArteryFlow Technology Corporation.

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. Hwang JH, Roh HG, Chun YI, Kang HS, Choi JW, Moon WJ, et al. Endovascular coil embolization of very small intracranial aneurysms. *Neuroradiology.* (2011) 53:349–57. doi: 10.1007/s00234-010-0735-0

2. Kwon BJ, Im SH, Park JC, Cho YD, Kang HS, Kim JE, et al. Shaping and navigating methods of microcatheters for endovascular treatment of paraclinoid aneurysms. *Neurosurgery.* (2010) 67:34–40. doi: 10.1227/01.NEU.0000370891.67129.2F

3. Kim TG. Optimal microcatheter shaping method customized for a patient-specific vessel using a three-dimensional printer. *J Cerebrovasc Endovasc Neurosurg.* (2021) 23:16–22. doi: 10.7461/jcen.2021.E2020.08.005

4. Namba K, Higaki A, Kaneko N, Mashiko T, Nemoto S, Watanabe E. Microcatheter shaping for intracranial aneurysm coiling using the 3-dimensional printing rapid prototyping technology: preliminary result in the first 10 consecutive cases. *World Neurosurg.* (2015) 84:178–86. doi: 10.1016/j.wneu.2015.03.006

5. Ishibashi T, Takao H, Suzuki T, Yuki I, Kaku S, Kan I, et al. Tailor-made shaping of microcatheters using three-dimensional printed vessel models for endovascular coil embolization. *Comput Biol Med.* (2016) 77:59–63. doi: 10.1016/j.compbiomed.2016.07.005

6. Xu Y, Tian W, Wei Z, Li Y, Dong B. Microcatheter shaping using three-dimensional printed models for intracranial aneurysm coiling. *J Neurointerv Surg.* (2020) 12:308–10. doi: 10.1136/neurintsurg-2019-015346

7. Nakajima S, Sakamoto M, Yoshioka H, Uno T, Kurosaki M. A new method of microcatheter heat-forming for cerebral aneurysmal coiling using stereolithography three-dimensional printed hollow vessel models. *Yonago Acta Med.* (2021) 64:113–9. doi: 10.33160/yam.2021.02.001

8. Ohshima T, Miyachi S, Matsuo N, Kawaguchi R, Takayasu M. A novel approach involving a microcatheter tip during aneurysmal coil embolization: the *Γ* (gamma) tip method. *Interv Neuroradiol.* (2019) 25:681–4. doi: 10.1177/1591019919850849

9. Liang XD, Wang ZL, Li TX, He YK, Bai WX, Wang YY, et al. Safety and efficacy of a new prophylactic tirofiban protocol without oral intraoperative antiplatelet therapy for endovascular treatment of ruptured intracranial aneurysms. *J Neurointerv Surg.* (2016) 8:1148–53. doi: 10.1136/neurintsurg-2015-012055

10. Shinoda N, Mori M, Tamura S, Korosue K, Kohmura E. Three-dimensional shaping technique for coil placement using the steam-shaped microcatheter for ruptured blood blister-like aneurysm. *Neurochirurgie.* (2018) 64:216–8. doi: 10.1016/j.neuchi.2018.04.006

11. Teresa NM, Jens F, Johanna S, Jan-Hendrik B, Maximilian FA. Current status of training environments in neuro-interventional practice: are animal models still contemporary? *J Neurointerv Surg.* (2019) 11:283–9. doi: 10.1136/neurintsurg-2018-014036

12. Liu CY, Shen Y, Wu XX, Qian K, Hu XB, Yang HF. Artificial intelligence-assisted microcatheter shaping for intracranial aneurysm coiling: a preliminary study. *Ann Vasc Surg.* (2022) 85:228–36. doi: 10.1016/j.avsg.2022.03.013

13. Ni wei. Microcatheter shaping for intracranial aneurysm coiling using artificial intelligence-assisted technique. *Neurosurgery.* (2019) 66:310–438. doi: 10.1093/neuros/nyz310_438

14. Hahnemann ML, Ringelstein A, Sandalcioglu IE, Goericke S, Moenninghoff C, Wanke I, et al. Silent embolism after stentassisted coiling of cerebral aneurysms: diffusion-weighted MRI study of 75 cases. *J Neurointerv Surg.* (2014) 6:461–5. doi: 10.1136/neurintsurg-2013-010820

15. Cagnazzo F, Ahmed R, Dargazanli C, Lefevre PH, Gascou G, Derraz I, et al. Treatment of wide-neck intracranial aneurysms with the Woven EndoBridge device associated with stenting: a singlecenter experience. *Am J Neuroradiol.* (2019) 40:820–6. doi: 10.3174/ajnr.A6032

16. Willinsky RA, Taylor SM, TerBrugge K, Farb RI, Tomlinson G, Montanera W. Neurologic complications of cerebral angiography: prospective analysis of 2,899 procedures and review of the literature. *Radiology.* (2003) 227:522–8. doi: 10.1148/radiol.2272012071

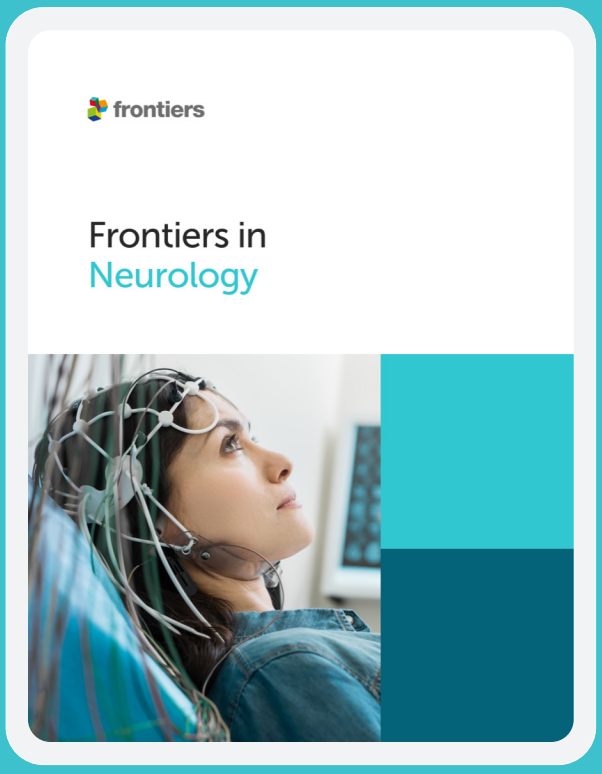
Frontiers in Neurology

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