Therapeutics and Clinical Risk Management

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Development and Validation of Machine Learning Models for Outcome Prediction in Patients with Poor-Grade Aneurysmal Subarachnoid Hemorrhage Following Endovascular Treatment

Senlin Du[1–4,*], Yanze Wu[1–4,*], Jiarong Tao[1], Lei Shu[1–4], Tengfeng Yan[1–4], Bing Xiao[1], Shigang Lv[1], Minhua Ye[1], Yanyan Gong[1], Xingen Zhu[1–4], Ping Hu[1,5], Miaojing Wu[1]

[1]Department of Neurosurgery, The second Affiliated Hospital, Jiangxi Medical College of Nanchang University, Nanchang, 330006, People's Republic of China; [2]Jiangxi Key Laboratory of Neurological Tumors and Cerebrovascular Diseases, Nanchang, 330006, People's Republic of China; [3]Jiangxi Health Commission Key Laboratory of Neurological Medicine, Nanchang, 330006, People's Republic of China; [4]Institute of Neuroscience, Nanchang University, Nanchang, 330006, People's Republic of China; [5]Department of Neurosurgery, Panzhihua Central Hospital, The second Clinical Medical College of Panzhihua University, Panzhihua, 617067, People's Republic of China

*These authors contributed equally to this work

Correspondence: Miaojing Wu, Department of Neurosurgery, The second Affiliated Hospital, Jiangxi Medical College of Nanchang University, Nanchang, 330006, People's Republic of China, Email wmj1987@163.com

Background: Endovascular treatment (EVT) has been recommended as a superior modality for the treatment of intracranial aneurysm. However, there still exists a worse percentage of poor functional outcome in patients with poor-grade aneurysmal subarachnoid hemorrhage (aSAH) undergoing EVT. Therefore, it is urgently needed to investigate the risk factors and develop a critical decision model in the subtype of such patients.

Methods: We extracted the target variables from an ongoing registry cohort study, PROSAH-MPC, which was conducted in multiple centers in China. We randomly assigned these patients to training and validation cohorts with a ratio of 7:3. Univariate and multivariate logistic regressions were performed to find the potential factors, and then nine machine learning models and a stack ensemble model were developed with optimized variables. The performance of these models was evaluated through several indicators, including area under the receiver operating characteristic curve (AUC-ROC). We further use Shapley Additive Explanations (SHAP) methods for the distribution of feature visualization based on the optimal models.

Results: A total of 226 eligible patients with poor-grade aSAH undergoing EVT were enrolled, while 89 (39.4%) has a poor 12-month outcome. Age (Adjusted OR [aOR], 1.08; 95% CI: 1.03–1.13; p = 0.002), subarachnoid hemorrhage volume (aOR, 1.02; 95% CI: 1.00–1.05; p = 0.033), World Federation of Neurosurgical Societies grade (WFNS) (aOR, 2.03; 95% CI: 1.05–3.93; p = 0.035), and Hunt-Hess grade (aOR, 2.36; 95% CI: 1.13–4.93; p = 0.022) were identified as the independent risk factors of the poor outcome. Then, the prediction models developed have revealed that LightGBM algorithm has a superior performance with an AUC-ROC value of 0.842 in the validation cohort, while the SHAP results showed that age is the most important risk factor affecting functional outcomes.

Conclusion: The LightGBM model holds immense potential in facilitating risk stratification for poor-grade aSAH patients undergoing endovascular treatment who are at risk of adverse outcomes, thereby enhancing clinical decision-making processes.

Trial Registration: PROSAH-MPC. NCT05738083. Registered 16 November 2022 – Retrospectively registered, https://clinicaltrials.gov/study/NCT05738083.

Keywords: intracranial aneurysm, subarachnoid hemorrhage, endovascular procedures, machine learning, prognosis

Introduction

Aneurysmal subarachnoid hemorrhage (aSAH) is a prevalent type of stroke associated with substantial mortality and morbidity rates, affecting approximately 30% of survivors with significant neurological impairments.[1] The International

Received: 14 November 2024
Accepted: 25 February 2025
Published: 7 March 2025

Subarachnoid Aneurysm Trial (ISAT)[2] has emphasized that endovascular coiling treatment is more likely to lead to independent survival at one year compared to neurosurgical clipping, with this survival benefit extending for at least seven years in cases where the ruptured aneurysm is suitable for both treatment options. However, a grim reality emerges when considering poor-grade aSAH, which classified as Hunt and Hess grades III and V, affects over 40% of SAH patients as our previous research indicates.[3] Despite significant advancements in endovascular treatment (EVT) and neurological intensive care, the prognosis for these patients remains extremely poor.[4–6]

A well-conducted systematic review has revealed a significant trend: the proportion of patients with poor-grade aSAH undergoing EVT have surged from 10.0% in the 1990–2000 period to an impressive 62.0% between 2010 and 2014. This shift has been accompanied by a gradual improvement in favorable neurological outcomes, increasing from 37.0% to 44.0% over the same timeframe. While an established model[7] with a superior performance can early predict the risk of poor outcomes in patients with aSAH receiving EVT, and even some models developed to assess the risk of outcome in the subtype of poor-grade aSAH,[8–10] these studies suffer from significant limitations. These include small sample sizes and the absence of external validation, which may lead to overfitting and limit the generalizability of their findings. Moreover, despite the superior performance of these models, poor interpretability and transparency hinder their clinical applicability.

Recently, machine learning (ML) algorithms have become powerful tools for analyzing complex medical datasets, surpassing traditional methods in predicting clinical outcomes.[11,12] However, poor-grade aSAH patients undergoing endovascular treatment (EVT) are often excluded from broader cohorts due to their severe condition. Existing models for this subgroup lack transparency, particularly in integrating Shapley Additive Explanations (SHAP), limiting their clinical applicability. Given the variability in prognosis and the complexity of EVT outcomes, real-time prognostic tools are essential for guiding individualized treatment. ML-based models, like the one proposed in this study, can aid in early risk identification and targeted interventions to improve outcomes and reduce complications such as delayed cerebral ischemia and rebleeding.

Hence, this study aims to develop a predictive model for poor outcomes in poor-grade aSAH patients undergoing EVT by applying advanced ML algorithms and newly measured data. We will compare model performance to identify the most effective approach and incorporate SHAP analysis to enhance interpretability.[13] Additionally, our model introduces a novel prognostic factor—total bleeding volume (TBV), identified in our previous research—distinguishing it from existing models. By focusing on this high-risk population, our study seeks to provide a clinically relevant and actionable framework for risk stratification and personalized treatment.

Methods and Materials

Study Design and Cohort

The PROSAH-MPC registry cohort study, identified by the number NCT05738083, is an investigator-initiated effort among multiple neurological centers in China. Its primary goal is to identify prognostic factors and establish robust prediction models that can accurately forecast complications, disability, and mortality in patients with aneurysmal SAH. Specifically, for this study, we have extracted data from eligible patients with poor-grade aSAH, classified as Hunt and Hess grades III and V, who underwent endovascular treatment (EVT) between October 2018 and December 2021. By focusing on this subset of patients, we aim to gain insights into the factors that influence their outcomes and develop predictive models tailored to their unique characteristics. The diagnosis of aneurysmal SAH in this study was rigorously confirmed using imaging modalities such as computed tomography (CT), CT angiography, or digital subtraction angiography (DSA), adhering strictly to the current guidelines. Furthermore, the study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) protocol.

The inclusion criteria defined as follows: (1) patients with spontaneous aneurysmal subarachnoid hemorrhage (aSAH); (2) patients with Hunt and Hess grades III and V at admission, indicating severe neurological impairment; (3) patients who received endovascular treatment (EVT) within 72 hours after the onset of symptoms; (4) patients who underwent a non-contrast computed tomography (CT) scan at the time of admission; and (5) patients who were available for follow-up for at least 1 year after discharge. Then we excluded patients with (1) patients with complicated cerebrovascular malformations or other pseudo-aneurysms; (2) patients with permanent brain injury at admission; (3) patients in a postoperative state at admission; (4) patients with incomplete clinical data. Figure 1 illustrates the detail of population enrollment from the dataset.

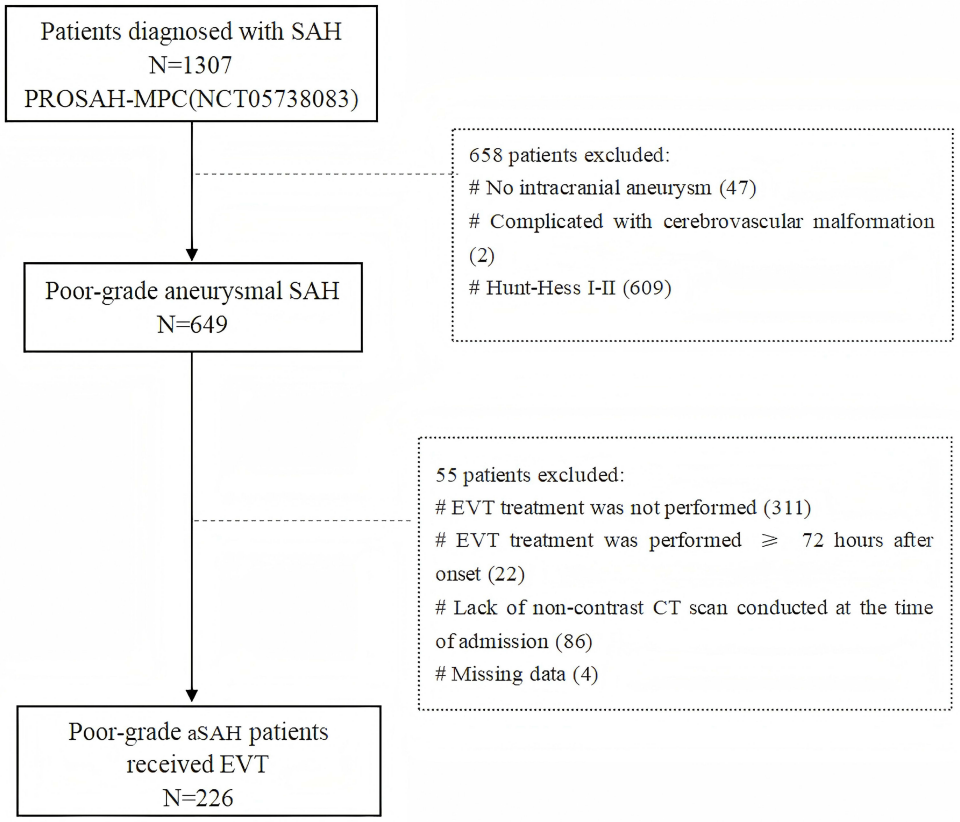

Figure 1 The schematic diagram illustrates the current research work and the corresponding abstract of the study.

Variables Collection

The target information was extracted from the Electronic Data Capture (EDC) database from the project of the PROSAH-MPC. This comprehensive dataset included a range of demographic information, such as age, sex, and relevant medical history factors that could potentially influence the patient's prognosis, including hypertension, diabetes, coronary heart disease, tobacco and alcohol consumption, and anticoagulant therapy. Additionally, the severity of the patients' condition on admission was assessed and extracted using several validated scales, including the World Federation of Neurosurgical Societies (WFNS) scale, the Hunt and Hess (HH) grade, and the modified Fisher scale (mFS). Detailed aneurysm features, including its location, number, length, width, and neck size, were also collected. We also extracted the condition of intracranial hemorrhage (total bleeding volume [TBV] and presence of intraparenchymal hemorrhage [IPH] and intraventricular hemorrhage [IVH]). It is noted that the total bleeding volume (TBV) was calculated using a proposed Hybrid 2D/3D U-Net model from our previous study.[14]

Missing Data Processing

Four patients had missing demographic data, representing less than 5% of the total patient population.[15,16] Consequently, the data were handled using a direct deletion approach.

Operation Management

All EVTs were performed by highly experienced senior neurointerventionalists. To prevent and manage cerebral vasospasm, all patients received intravenous nimodipine for up to 21 days postoperatively, in accordance with current clinical guidelines. Nimodipine is widely recognized for its efficacy in reducing the risk of delayed cerebral ischemia associated with vasospasm, as recommended by the American Heart Association (AHA) guidelines.[1,17]

https://doi.org/10.2147/TCRM.S504745

For the treatment of brain edema, osmotic therapy was administered using either mannitol or hypertonic saline, depending on intracranial pressure levels. This strategy aligns with established protocols for controlling elevated intracranial pressure and mitigating the risk of further neurological deterioration.[1]

Outcome Definition

The neurological outcomes of these patients were evaluated at the 12-month mark following the initial stroke onset, utilizing the modified Rankin Scale (mRS) as the assessment tool. A favorable neurological outcome was designated as an mRS score within the range of 0 to 2, indicative of minimal to no disability. Conversely, a poor outcome was classified as an mRS score spanning from 3 to 6, suggesting moderate-to-severe disability or even death. To ensure objectivity, all patient follow-ups were conducted via telephone consultations with a neurosurgeon who was blinded to the patients' clinical information.

Model Development and Validation

Initially, the Boruta algorithm was employed to identify the most pivotal factors influencing the outcomes of endovascularly treated aSAH patients. Leveraging the inherent stability and credibility of the multiple random forest classification algorithm, the Boruta algorithm successfully extracted robust and reliable features.

Subsequently, the dataset was meticulously divided into training and validation sets, maintaining a ratio of 7:3. The training set served as the foundation for constructing nine distinct machine learning (ML) models, encompassing Logistic Regression (LR), Decision Tree (DT), Elastic Net (Enet), K-Nearest Neighbors (KNN), Light Gradient Boosting Machine (LightGBM), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machines (SVM), and Multilayer Perceptron (MLP). We used a grid search approach combined with five-fold cross-validation to determine the optimal hyperparameters for each model. Detailed Hyperparameters can be founded in Table S1.

To enhance predictive performance, we ventured beyond individual models and employed the Least Absolute Shrinkage and Selection Operator (LASSO) to develop a sophisticated stacking ensemble model. This ensemble model adeptly fused the insights from the nine individual classifiers, offering a comprehensive and integrated perspective.

Logistic regression was chosen as the benchmark model for comparative analysis, owing to its widespread adoption in prior medical research for linear predictive tasks. Our primary objective was to assess whether a non-linear ML approach could offer incremental benefits, surpassing the performance of the traditional linear model.

Model Interpretability

The SHAP algorithm is adopted to elucidate the reliability and importance of model predictions. By assigning each variable its corresponding attribution value (SHAP value), which can be used to quantitatively measure the impact of each feature and sample on the model predictions and thus interpret prediction results. The SHAP summary plot was employed to illustrate the contributions of each feature attributed to the model. Moreover, the SHAP force plot was further used to visualize the impact of crucial features on the final model for individual patients.

Statistical Analysis

The Kolmogorov–Smirnov test served as the cornerstone in identifying the nature of variable distributions. For continuous variables, we employed either the independent *t*-test or the Mann–Whitney *U*-test, presenting the results as Mean ± SD or median alongside the inter-quartile range (IQR), respectively. Categorical variables, on the other hand, were scrutinized using Chi-square or Fisher's exact tests, with outcomes expressed as percentages.

To ensure the pinnacle of optimization and robustness for each ML model, we integrated hyperparameter tuning with a rigorous five-fold cross-validation procedure. The validation group's model performance was meticulously evaluated using two pivotal metrics: AUROC and PRAUC. These metrics served as a sieve, enabling us to single out the optimal model that boasted the utmost predictive accuracy.

To gauge the calibration of our chosen model, we harnessed the Hosmer–Lemeshow goodness-of-fit test, which furnished a statistical gauge of how seamlessly the model's predictions aligned with actual outcomes. Furthermore, to

delve into the clinical significance of these algorithms, we conducted decision curve analysis (DCA). DCA provided a quantitative lens to assess the tangible benefits derived from incorporating a specific model into clinical decision-making, thereby illuminating the real-world implications of our discoveries.

For the SHAP value analysis, we leveraged the "fastshap" package within R software, while visualization of these values for each feature was masterfully achieved through the "ggbeeswarm" and "shapviz" packages. All statistical tests adhered to a two-tailed approach, with statistical significance set at $P < 0.05$. The statistical analyses were meticulously performed using IBM SPSS Statistics for Windows (Version 26.0, IBM Corp., Armonk, NY, USA) and R software (version 4.3.0, accessible at https://www.r-project.org/).

Results

Baseline Characteristics

A total of 226 patients with poor-grade aSAH receiving EVT were included in the study. We divided them into a training and validation cohorts according to the ratio of 7:3. There was no significant difference in baseline characteristics between the training and validation cohorts ($p>0.05$) (Table 1). The number of patients with poor outcome was 66 (41.8%) for the training cohort and 23 (33.8%) for the validation cohort. Among them, women represented 149 patients, accounting for 65.9% of the total. And the average age was 63.5 years. Table 2 showed the detailed baseline characteristics of the training and validation cohorts.

Table 1 Baseline Characteristics Between Training Set and Validation Set in Aneurysmal Subarachnoid Hemorrhage Following Endovascular Treatment

Characteristic	Overall	Training Set	Validation Set	P
Participants, No.	**226**	**158**	**68**	
Age (median [IQR])	63.50 [53.00, 69.75]	64.00 [54.00, 70.00]	62.50 [52.00, 69.00]	0.459
Female (%)	149 (65.9)	100 (63.3)	49 (72.1)	0.262
Hypertension (%)	131 (58.0)	90 (57.0)	41 (60.3)	0.750
Diabetes (%)	12 (5.3)	7 (4.4)	5 (7.4)	0.281
CHD (%)	8 (3.5)	5 (3.2)	3 (4.4)	0.942
Smoking (%)	22 (9.7)	16 (10.1)	6 (8.8)	0.953
Drinking (%)	9 (4.0)	8 (5.1)	1 (1.5)	0.370
Anticoagulant (%)	10 (4.4)	6 (3.8)	4 (5.9)	0.729
TBV (median [IQR])	29.67 [16.64, 49.71]	28.88 [17.20, 47.00]	31.32 [16.54, 53.86]	0.682
IPH (%)	173 (76.5)	121 (76.6)	52 (76.5)	>0.99
IVH (%)	64 (28.3)	42 (26.6)	22 (32.4)	0.470
WFNS (%)				0.292
II	61 (27.0)	40 (25.3)	21 (30.9)	
III	7 (3.1)	6 (3.8)	1 (1.5)	
IV	97 (42.9)	73 (46.2)	24 (35.3)	
V	61 (27.0)	39 (24.7)	22 (32.4)	
Hunt-Hess (%)				0.272
III	119 (52.7)	88 (55.7)	31 (45.6)	
IV	46 (20.4)	32 (20.3)	14 (20.6)	
V	61 (27.0)	38 (24.1)	23 (33.8)	
mFS (%)				0.937
1	5 (2.2)	4 (2.5)	1 (1.5)	
2	38 (16.8)	27 (17.1)	11 (16.2)	
3	82 (36.3)	58 (36.7)	24 (35.3)	
4	101 (44.7)	69 (43.7)	32 (47.1)	

(*Continued*)

Table 1 (Continued).

Characteristic	Overall	Training Set	Validation Set	P
Participants, No.	**226**	**158**	**68**	
Aneurysm multiple (%)	33 (14.6)	28 (17.7)	5 (7.4)	0.082
Size (median [IQR])				
Aneurysm Length	4.80 [3.40, 6.47]	4.80 [3.30, 6.60]	4.80 [3.50, 6.40]	0.910
Aneurysm Width	3.90 [3.00, 5.30]	3.95 [3.00, 5.47]	3.80 [3.00, 4.80]	0.507
Neck	3.15 [2.50, 4.10]	3.15 [2.50, 4.00]	3.15 [2.60, 4.43]	0.523
Location (%)				0.307
ACA	6 (2.7)	5 (3.2)	1 (1.5)	
MCA	22 (9.7)	19 (12.0)	3 (4.4)	
ACoA	76 (33.6)	55 (34.8)	21 (30.9)	
PCA	8 (3.5)	6 (3.8)	2 (2.9)	
PCoA	82 (36.3)	54 (34.2)	28 (41.2)	
Other	32 (14.2)	9 (12.0)	13 (19.1)	
Poor outcome	89 (39.4)	66 (41.8)	23 (33.8)	0.330

Abbreviations: WFNS, World Federation of Neurosurgical Societies; mFS, modified Fisher scale; ACA, anterior cerebral aneurysm; MCA, middle cerebral aneurysm; ACoA, anterior communicating aneurysm; PCA, posterior cerebral aneurysm; PCoA, posterior communicating aneurysm; TBV, total bleeding volume; IPH, intraparenchymal hemorrhage; IVH, intraventricular hemorrhage.

Univariate and Multivariate Logistic Regression

Univariate and multivariate logistic regression were employed to differentiate characteristics between outcomes, showing together distinguished factors associated with outcome at 12 months after discharge (Table 3). Age (Adjusted OR [aOR], 1.08; 95% CI: 1.03–1.13; p = 0.002), TBV (aOR, 1.02; 95% CI: 1.00–1.05; p = 0.033), Hunt-Hess grade (aOR, 2.36; 95% CI: 1.13–4.93; p = 0.022), and WFNS grade (aOR, 2.03; 95% CI: 1.05–3.93; p = 0.035) were deemed as contributing factors for poor outcome. The corresponding results of univariate and multivariate logistic regression were illustrated by a forest plot (Figure 2).

Model Performance

Potential predictive variables were split out using shadow features through the Boruta algorithm. Six most relevant features were employed to train and build the ensemble model, including the mFS, hypertension, TBV, age, Hunt-Hess and WFNS grade (Figure 3). In the training set, fivefold cross-validation was used to evaluate predictive performance and general error estimates in the model development. Next, we assessed the predictive capabilities of machine learning models that were trained using a combination of 9 distinct algorithms and a stacking ensemble model.

In the training and validation set, LightGBM exhibited superior predictive performance with an AUC-ROC values of 0.901 and 0.842, respectively (Figure 4A and B). The PR curve results indicated that the PRAUC values for the LightGBM model were distinguished, with corresponding values of 0.874 (training set) and 0.745 (validation set) (Figure 4C and D). Then, DCA was used to evaluate the clinical application value of each prediction model. As shown in Figure 4E and F, LightGBM model still exhibited the continuous maximum benefit in the training and validation set. The calibration curve showed a strong correlation between the predicted and actual risks in terms of Brier score (BS), which was used for indicating the calibration ability. The LightGBM model had the best calibration in the training group and validation group (Figure 4G and H). Table 4 records the details of each model performance for training and validation cohorts. Furthermore, we performed a comparative analysis between the LightGBM model and well-established clinical tools, including the Hunt-Hess and WFNS grading systems. As depicted in Figure S1, the area under the curve (AUC) for WFNS is 0.738 (95% CI: 0.682–0.794), whereas for the Hunt and Hess grading, it is 0.742 (95% CI: 0.691–0.793). This comparison highlights the superior predictive performance of our model over the traditional scoring systems.

Table 2 Patient Characteristics and Group Comparisons for Aneurysmal Subarachnoid Hemorrhage Following Endovascular Treatment Between Training Set and Validation Set

Characteristic	Training Set				Validation Set			
	Overall	Good Outcome	Poor Outcome	P	Overall	Good Outcome	Poor Outcome	P
Participants, No.	158	92	66		68	45	23	
Age (median [IQR])	64.00 [54.00, 70.00]	61.00 [52.00, 66.00]	68.00 [61.00, 73.00]	<0.001	62.50 [52.00, 69.00]	61.00 [50.00, 66.00]	69.00 [58.00, 73.50]	0.011
Female (%)	100 (63.3)	60 (65.2)	40 (60.6)	0.670	49 (72.1)	33 (73.3)	16 (69.6)	0.966
Hypertension (%)	90 (57.0)	46 (50.0)	44 (66.7)	0.054	41 (60.3)	25 (55.6)	16 (69.6)	0.392
Diabetes (%)	7 (4.4)	4 (4.3)	3 (4.5)	>0.99	5 (7.4)	1 (2.2)	4 (17.4)	0.076
CHD (%)	5 (3.2)	3 (3.3)	2 (3.0)	>0.99	3 (4.4)	3 (6.7)	0 (0.0)	0.521
Smoking (%)	16 (10.1)	10 (10.9)	6 (9.1)	0.922	6 (8.8)	6 (13.3)	0 (0.0)	0.167
Drinking (%)	8 (5.1)	5 (5.4)	3 (4.5)	>0.99	1 (1.5)	1 (2.2)	0 (0.0)	>0.99
Anticoagulant (%)	6 (3.8)	3 (3.3)	3 (4.5)	>0.99	4 (5.9)	3 (6.7)	1 (4.3)	>0.99
TBV (median [IQR])	28.88 [17.20, 47.00]	25.24 [13.37, 37.52]	38.58 [26.55, 60.36]	<0.001	31.32 [16.54, 53.86]	25.97 [13.19, 34.63]	56.29 [34.72, 83.36]	<0.001
IPH (%)	121 (76.6)	67 (72.8)	54 (81.8)	0.260	52 (76.5)	30 (66.7)	22 (95.7)	0.018
IVH (%)	42 (26.6)	19 (20.7)	23 (34.8)	0.070	22 (32.4)	15 (33.3)	7 (30.4)	>0.99
WFNS				<0.001				0.006
II	40 (25.3)	35 (38.0)	5 (7.6)		21 (30.9)	16 (35.6)	5 (21.7)	
III	6 (3.8)	4 (4.3)	2 (3.0)		1 (1.5)	0 (0.0)	1 (4.3)	
IV	73 (46.2)	47 (51.1)	26 (39.4)		24 (35.3)	20 (44.4)	4 (17.4)	
V	39 (24.7)	6 (6.5)	33 (50.0)		22 (32.4)	9 (20.0)	13 (56.5)	
Hunt-Hess				<0.001				0.015
III	88 (55.7)	68 (73.9)	20 (30.3)		31 (45.6)	25 (55.6)	6 (26.1)	
IV	32 (20.3)	18 (19.6)	14 (21.2)		14 (20.6)	10 (22.2)	4 (17.4)	
V	38 (24.1)	6 (6.5)	32 (48.5)		23 (33.8)	10 (22.2)	13 (56.5)	
mFS				0.006				0.177
1	4 (2.5)	2 (2.2)	2 (3.0)		1 (1.5)	0 (0.0)	0 (0.0)	
2	27 (17.1)	21 (22.8)	6 (9.1)		11 (16.2)	10 (22.2)	1 (4.3)	
3	58 (36.7)	39 (42.4)	19 (28.8)		24 (35.3)	16 (35.6)	8 (34.8)	
4	69 (43.7)	30 (32.6)	39 (59.1)		32 (47.1)	18 (40.0)	14 (60.9)	
Aneurysm multiple (%)	28 (17.7)	12 (13.0)	16 (24.2)	0.108	5 (7.4)	3 (6.7)	2 (8.7)	>0.99
Size (median [IQR])								
Length	4.80 [3.30, 6.60]	4.50 [2.98, 6.12]	5.10 [4.03, 6.77]	0.092	4.80 [3.50, 6.40]	4.80 [3.40, 6.20]	5.00 [3.55, 6.85]	0.433
Width	3.95 [3.00, 5.47]	3.90 [2.92, 5.05]	4.05 [3.00, 5.77]	0.434	3.80 [3.00, 4.80]	3.80 [3.00, 4.80]	3.80 [3.00, 5.00]	0.731
Neck	3.15 [2.50, 4.00]	3.08 [2.28, 4.00]	3.50 [2.60, 4.07]	0.253	3.15 [2.60, 4.43]	3.40 [2.60, 4.50]	3.10 [2.70, 3.85]	0.683
Location (%)				0.859				0.429
ACA	5 (3.2)	3 (3.3)	2 (3.0)		1 (1.5)	1 (2.2)	0 (0.0)	
MCA	19 (12.0)	9 (9.8)	10 (15.2)		3 (4.4)	2 (4.4)	1 (4.3)	
ACoA	55 (34.8)	32 (34.8)	23 (34.8)		21 (30.9)	14 (31.1)	7 (30.4)	
PCA	6 (3.8)	3 (3.3)	3 (4.5)		2 (2.9)	0 (0.0)	2 (8.7)	
PCoA	54 (34.2)	32 (34.8)	22 (33.3)		28 (41.2)	20 (44.4)	8 (34.8)	
Other	19 (12.0)	13 (14.1)	6 (9.1)		13 (19.1)	8 (17.8)	5 (21.7)	

Abbreviations: WFNS, World Federation of Neurosurgical Societies; mFS, modified Fisher scale; ACA, anterior cerebral aneurysm; MCA, middle cerebral aneurysm; ACoA, anterior communicating aneurysm; PCA, posterior cerebral aneurysm; PCoA, posterior communicating aneurysm; TBV, total bleeding volume; IPH, intraparenchymal hemorrhage; IVH, intraventricular hemorrhage; PSM, propensity score matching.

https://doi.org/10.2147/TCRM.S504745

Table 3 Association Between Treatment Modality and Functional Outcome in Univariate and Multivariate Logistic Regression Analysis

Characteristic	Univariate Analysis		Multivariate Analysis	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Female	0.82 (0.43–1.58)	0.553		
Age	1.07 (1.03–1.11)	<0.001	1.08 (1.03–1.13)	0.002
Hypertension	2.00 (1.04–3.85)	0.038	2.39 (0.99–5.80)	0.053
Diabetes	1.05 (0.23–4.85)	0.953		
CHD	0.93 (0.15–5.71)	0.935		
Smoking	0.82 (0.28–2.38)	0.715		
Drinking	0.83 (0.19–3.60)	0.802		
Anticoagulant	1.41 (0.28–7.23)	0.678		
TBV	1.03 (1.01–1.04)	<0.001	1.02 (1.00–1.05)	0.033
IPH	1.68 (0.77–3.65)	0.191		
IVH	2.06 (1.01–4.20)	0.048	1.60 (0.61–4.17)	0.337
WFNS	3.18 (2.05–4.93)	<0.001	2.03 (1.05–3.93)	0.035
Hunt-Hess	3.95 (2.49–6.29)	<0.001	2.36 (1.13–4.93)	0.022
mFS	1.87 (1.22–2.88)	0.004	0.60 (0.29–1.25)	0.173
Aneurysm length	1.03 (0.93–1.14)	0.572		
Aneurysm width	0.99 (0.89–1.11)	0.923		
Aneurysm neck width	0.96 (0.85–1.09)	0.543		
Location				
ACA	Ref			
MCA	1.67 (0.22–12.35)	0.617		
ACoA	1.08 (0.17–6.98)	0.937		
PCA	1.50 (0.14–16.54)	0.741		
PCoA	1.03 (0.16–6.69)	0.974		
Other	0.69 (0.09–5.29)	0.723		

Abbreviations: WFNS, World Federation of Neurosurgical Societies; mFS, modified Fisher scale; TBV, total bleeding volume; IPH, intraparenchymal hemorrhage; IVH, intraventricular hemorrhage.

In general, it is evident that the LightGBM model exhibited superior performance compared to other models, and there was no evidence of overfitting in both the training and validation sets. Therefore, for subsequent analysis, the interpretability of the optimal model (LightGBM) was prioritized.

Model Interpretability

The SHAP analysis was conducted to assess the significance of features in the LightGBM model, considering their global importance and specific classification outcomes. These findings are illustrated in Figure 5A and B. The feature importance ranking for developing poor outcome is as follows: age, WFNS, Hunt-Hess grade, TBV, hypertension and mFS scale.

To enhance the understanding of the variables in the predictive model, the SHAP dependency plot for all six features was generated (Figure 5C and D). Old age, high Hunt-Hess grade, high WFNS grade, elevated TBV levels, high modified Fisher scale (mFS) grade, and a history of hypertension were all associated with an increased risk of poor outcomes. Furthermore, the effects of age and TBV levels on poor outcomes represent a non-linear pattern. We found that being older than 60 years significantly increased the risk of poor outcome and that TBV values higher than ~25 mL were strong indicators of an increased the risk of adverse outcome (Figure 5D).

The SHAP force plot (Figure 6) illustrates model interpretation at the individual level (Figure 6). Figure 6A depicts a low-risk patient, in which the patient was 63 years old, had high Hunt-Hess, WFNS and mFS scores that collectively contributed negatively to their poor prognosis. Additionally, hypertension increased the patient's risk of a poor prognosis. Figure 6B presents risk prediction process for a high-risk patient, primarily driven by older age and a larger TBV. However, a low Hunt-Hess grade was found to be a protective factor for prognosis, while WFNS and mFS had a weak positive association with poor outcome.

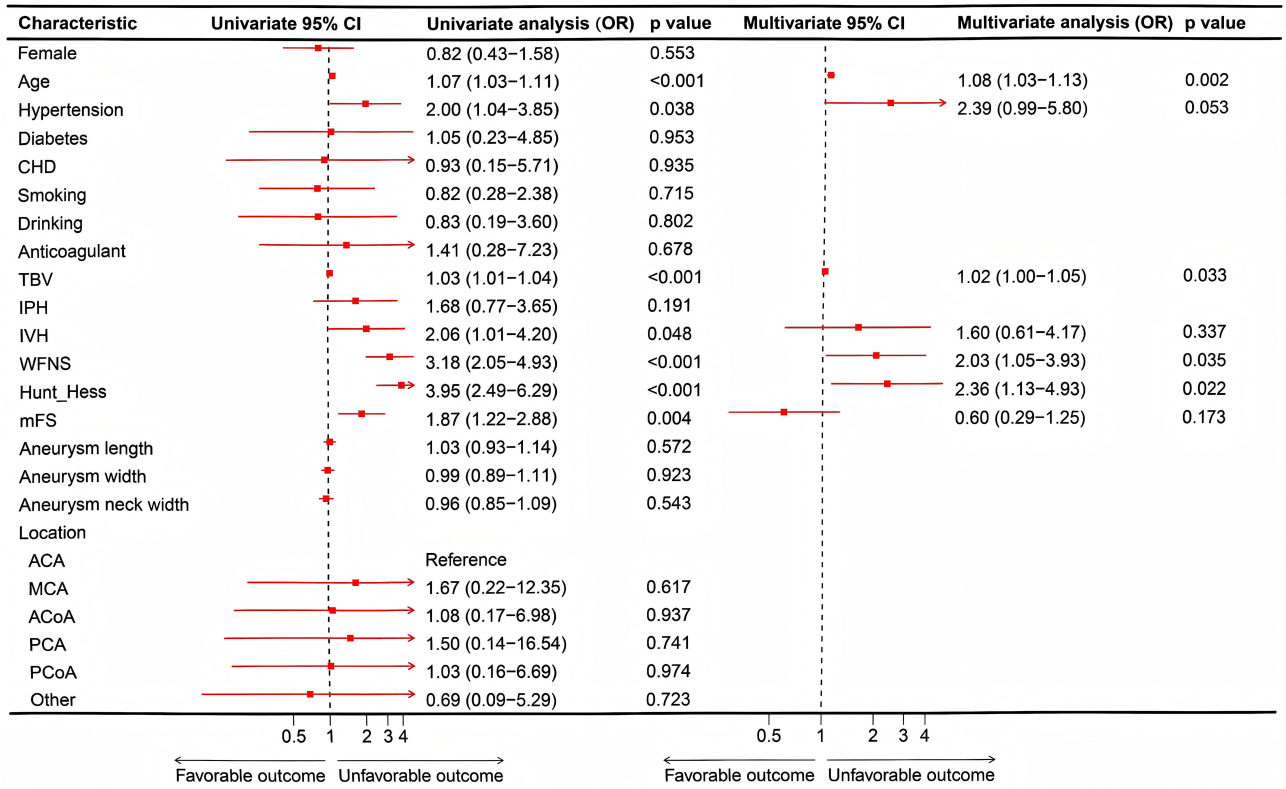

Figure 2 The forest plot of univariate and multivariate logistic regression analyses for poor outcome in high-grade aSAH following endovascular treatment.
Abbreviations: CHD, Coronary Heart Disease; TBV, Total Bleeding Volume; IPH, Intraparenchymal Hemorrhage; IVH, Intraventricular Hemorrhage; WFNS, World Federation of Neurological Societies; Hunt-Hess, Hunt and Hess grade; mFS, Modified Fisher Scale; ACA, Anterior Cerebral Artery; MCA, Middle Cerebral Artery; ACoA, Anterior Communicating Artery; PCA, Posterior Cerebral Artery; PCoA, Posterior Communicating Artery.

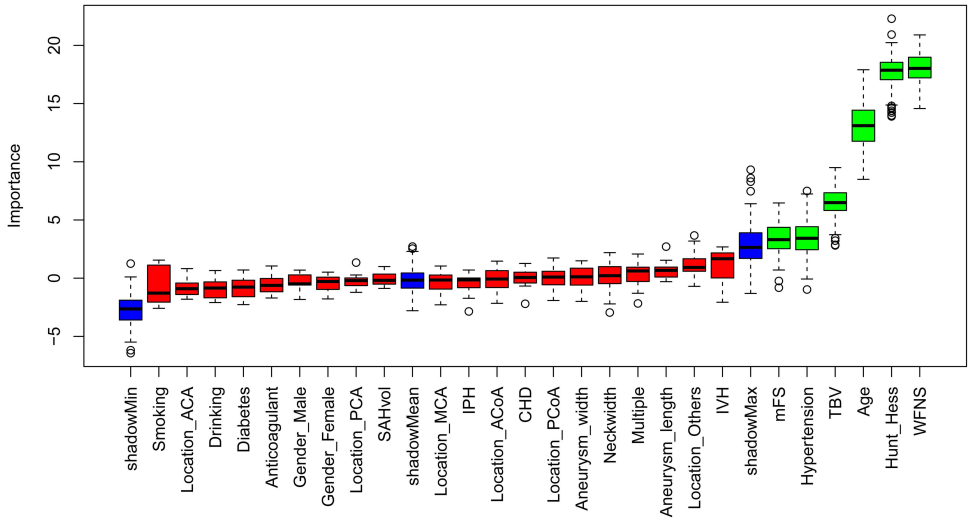

Figure 3 Feature selection technique: Boruta result plot for training data. Blue boxplots correspond to the minimal, average, and maximum *Z* scores of shadow attributes. Red boxplots represent the *Z* scores of rejected attributes, while green boxplots represent the *Z* scores of confirmed attributes.
Abbreviations: ACA, Anterior Cerebral Artery; PCA, Posterior Cerebral Artery; SAHvol, Subarachnoid Hemorrhage Volume; MCA, Middle Cerebral Artery; IPH, Intraparenchymal Hemorrhage; ACoA, Anterior Communicating Artery; CHD, Coronary Heart Disease; PCoA, Posterior Communicating Artery; IVH, Intraventricular Hemorrhage; mFS, Modified Fisher Scale; TBV, Total Bleeding Volume; Hunt-Hess, Hunt and Hess grade; WFNS, World Federation of Neurological Societies.

https://doi.org/10.2147/TCRM.S504745

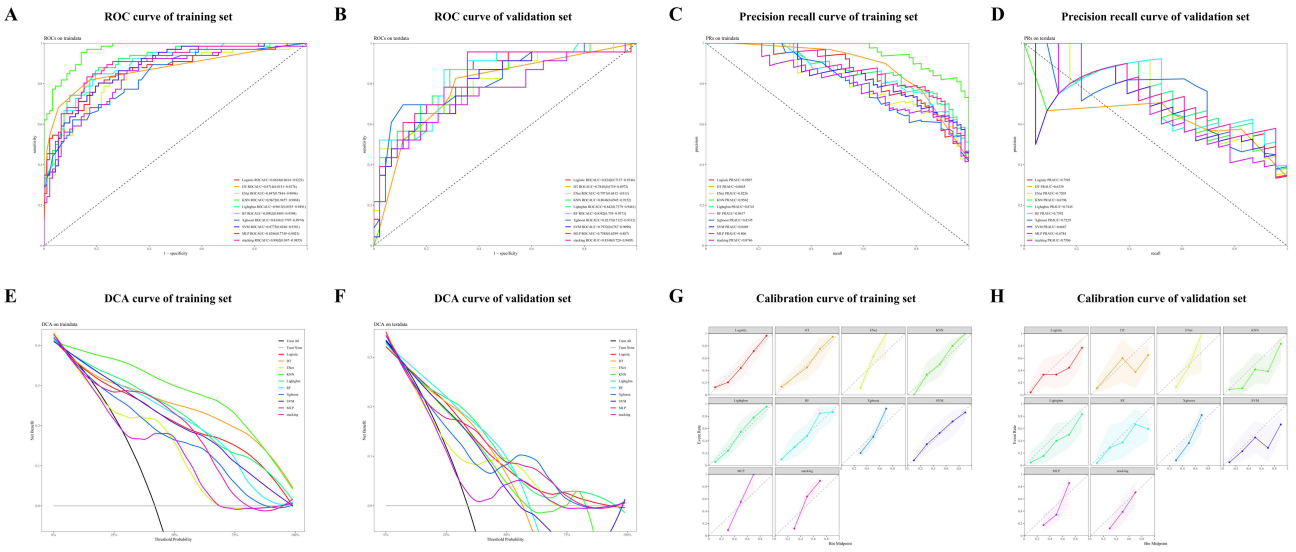

Figure 4 Performance of the models in training set (**A, C, E, G**) and validation set (**B, D, F, H**). (**A**) The ROC curve of each model in the training set; (**B**) The ROC curve of each model in the validation set; (**C**) The precision-recall of each model in the training set; (**D**) The precision-recall of each model in the validation set; (**E**) The DCA curve of each model in the training set; (**F**) The DCA curve of each model in the validation set; (**G**) The calibration curve of each model in the training set; (**H**) The calibration curve of each model in the validation set.

Abbreviations: ROC, Receiver Operating Characteristic; DCA, Decision curve analysis.

Discussion

In this study, we trained nine ML models and a stacking model specifically tailored to analyze the dataset of poor-grade aSAH patients undergoing EVT. Notably, the LightGBM model emerged as the most clinically predictive, achieving remarkable AUROC and PRAUC scores of 0.842 and 0.7445, respectively, in the validation set. To enhance both the model's effectiveness and interpretability, we integrated the SHAP technique, providing deeper insights into its decision-

Table 4 Model Performance Using Training and Validation Cohorts

Cohort	Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	Precision	Recall
Training	LR	0.930 (0.880–0.979)	0.877	0.848	0.893	0.813	0.848
	DT	0.880 (0.815–0.946)	0.900	0.761	0.976	0.946	0.761
	ENet	0.903 (0.845–0.962)	0.877	0.783	0.929	0.857	0.783
	KNN	0.993 (0.985–1.000)	0.969	0.957	0.976	0.957	0.957
	Lightgbm	0.905 (0.846–0.964)	0.877	0.804	0.917	0.841	0.804
	RF	0.904 (0.842–0.967)	0.892	0.826	0.929	0.864	0.826
	Xgboost	0.897 (0.830–0.964)	0.869	0.848	0.881	0.796	0.848
	SVM	0.916 (0.860–0.973)	0.885	0.826	0.917	0.844	0.826
	MLP	0.901 (0.845–0.957)	0.877	0.761	0.940	0.875	0.761
	Stacking						
Validation	LR	0.781 (0.657–0.905)	0.719	0.679	0.759	0.731	0.679
	DT	0.804 (0.695–0.914)	0.772	0.643	0.897	0.857	0.643
	ENet	0.866 (0.771–0.961)	0.772	0.643	0.897	0.857	0.643
	KNN	0.807 (0.688–0.925)	0.719	0.643	0.793	0.750	0.643
	Lightgbm	0.805 (0.688–0.923)	0.737	0.679	0.793	0.760	0.679
	RF	0.865 (0.771–0.960)	0.789	0.714	0.862	0.833	0.714
	Xgboost	0.882 (0.799–0.966)	0.737	0.714	0.758	0.741	0.714
	SVM	0.862 (0.768–0.957)	0.754	0.750	0.759	0.750	0.750
	MLP	0.853 (0.752–0.955)	0.754	0.679	0.828	0.792	0.679
	Stacking						

Abbreviations: LR, logistic regression; DT, decision tree; Enet, Elastic Networks; KNN, K Nearest Neighbors; Lightgbm, Light Gradient boosting machine; RF, random forest; Xgboost, eXtreme Gradient Boosting; SVM, support vector machine; MLP, Multilayer perceptron.

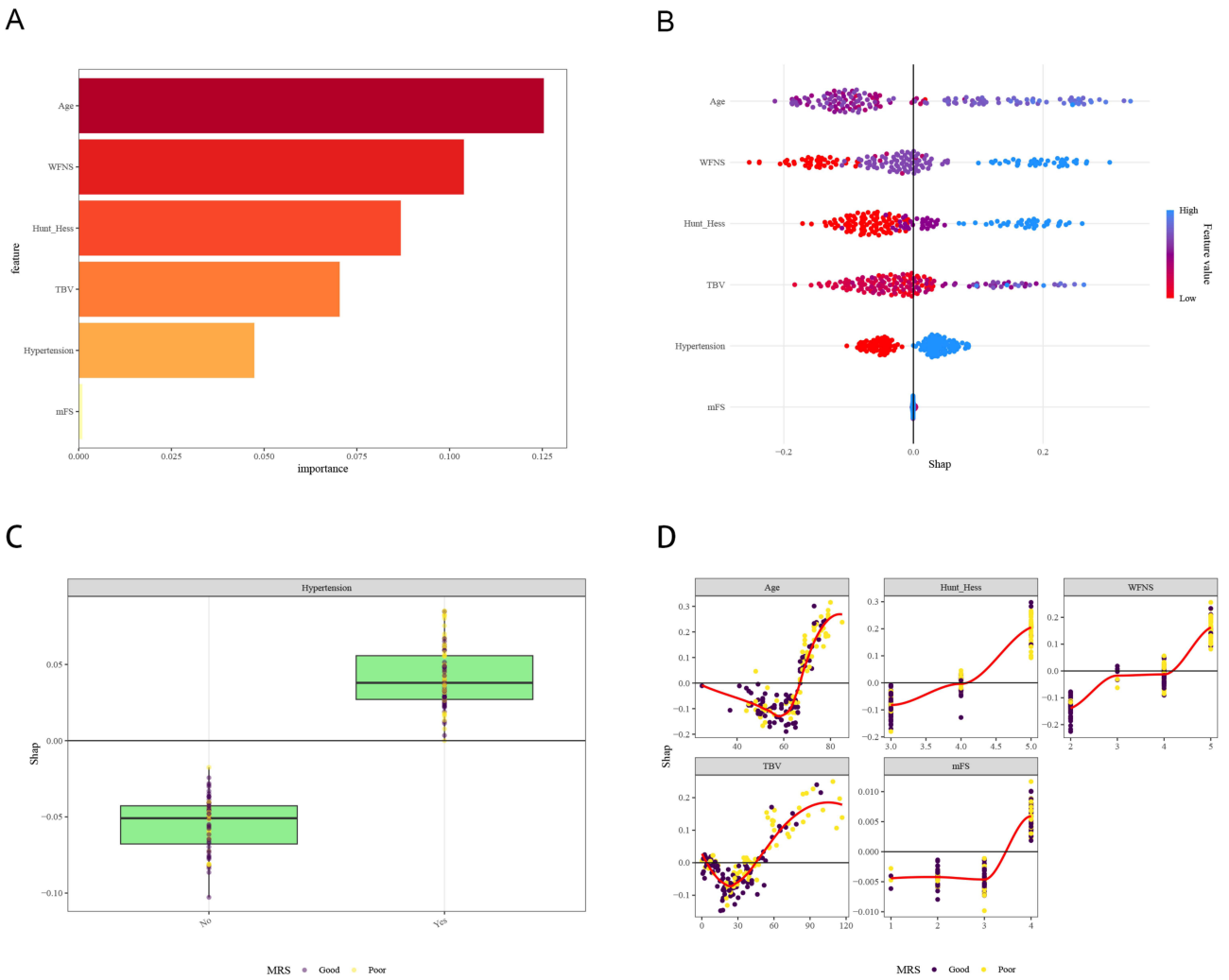

Figure 5 SHAP analysis of feature importance. (**A**) Feature importance ranking based on LightGBM; (**B**) Feature importance ranking based on SHAP values. (**C**) Dependence plot of categorical variables based on SHAP values; (**D**) Dependence plot of numerical variables based on SHAP values. The vertical axis lists features from top to bottom in order of decreasing importance. The position of a point on the horizontal axis indicates the feature's influence on the model's predicted value, while the point's color reflects the feature's value. For numerical variables, blue and red points represent higher and lower values, respectively; for categorical variables, blue and red points correspond to "yes" and "no", respectively.

Abbreviations: Hunt-Hess, Hunt and Hess grade; WFNS, World Federation of Neurological Societies; mFS, Modified Fisher Scale; TBV, Total Bleeding Volume.

making process. This integration is poised to significantly empower clinicians with a profound comprehension of the model's underlying reasoning, facilitating more informed and efficient utilization of its predictive insights in clinical practice.

Despite the persistent challenges of poor-grade aSAH, emerging evidence offers a promising outlook. Henry et al[18] underscores the transformative impact of well-informed clinical decisions on enhancing survivors' quality of life. Meanwhile, the growing recognition of EVT as a beneficial intervention for poor-grade aSAH patients, as evidenced by Ishikawa et al[19] underscores the pressing need for accurate long-term outcome predictions and risk factor identification. Recent advancements in predictive modeling have been remarkable. Liu et al[20] demonstrated that a decision tree model achieved an impressive AUC of 0.88 in predicting the prognosis of high-grade aSAH patients, while a novel scoring system[21] demonstrated heightened predictive accuracy, with an AUC of 0.831 in the validation cohort. ML models have outperformed traditional predictive models, yet their clinical applicability remains hindered by a lack of interpretability.[22] The integration of explainable ML has shown remarkable success across various medical domains,[23–26] highlighting its potential to bridge the gap between cutting-edge technology and clinical practice.

https://doi.org/10.2147/TCRM.S504745

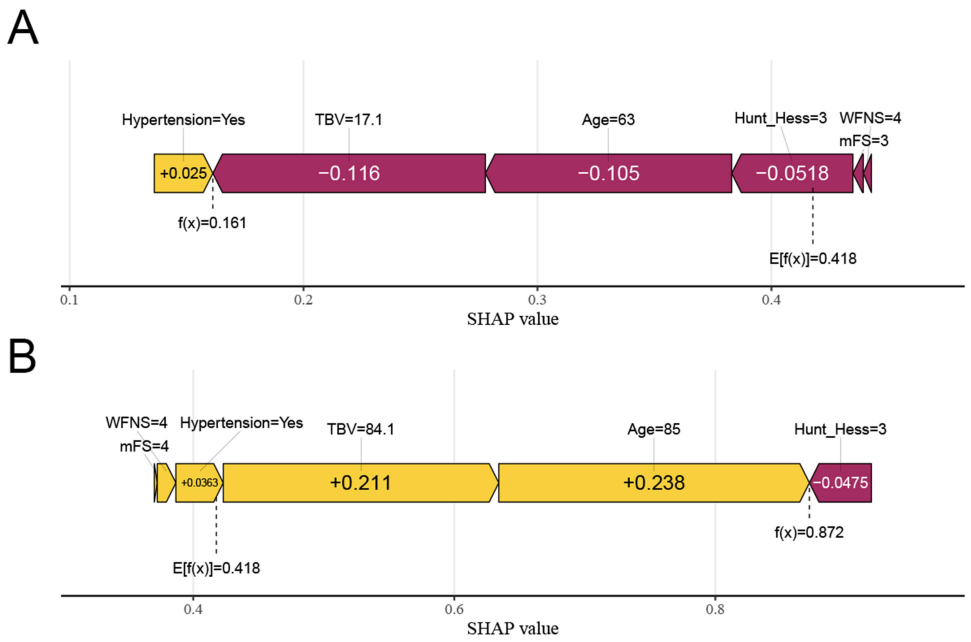

Figure 6 Specific prediction and interpretation of the lightGBM model for two patients. This plot offers a visual illustration of the LightGBM model's predictions, wherein the yellow and purple bars signify risk factors and protective factors, respectively. The length of the bars corresponds to the extent of feature importance. (**A**) Favorable outcome; (**B**) Poor outcome.

Abbreviations: Hunt-Hess, Hunt and Hess grade; WFNS, World Federation of Neurological Societies; mFS, Modified Fisher Scale; TBV, Total Bleeding Volume.

In this context, the introduction of SHAP analysis represents a significant advance, providing a game-theoretic approach that sheds light on the previously inscrutable "black box" of ML models.[12] To the best of our knowledge, this study is the first to employ SHAP analysis for predicting long-term outcomes in high-grade aSAH patients undergoing EVT, thereby enhancing both the interpretability and clinical applicability of ML in this critical domain. Our research conducted a comprehensive evaluation of multiple ML algorithms, ultimately identifying LightGBM as the most accuracy predictive model. LightGBM, a sophisticated ensemble of decision trees tailored for both classification and regression tasks, boasts widespread adoption across predictive modeling landscapes and holds significant practical implications.[27,28] However, recognizing its inherent black-box nature, we innovatively employed the SHAP methodology to reveal both global and local insights into the model's decision-making process.

The SHAP analysis highlighted the pivotal role of crucial clinical factors in predicting the long-term prognosis of poor-grade aSAH patients undergoing EVT. Age emerged as the predominant predictor, with SHAP analysis revealing a substantial increase in risk for patients over 60 years old. This finding aligns with existing literature, emphasizing age as a fundamental determinant of cerebrovascular prognosis.[29,30] Older patients, often burdened by comorbidities and reduced physiological resilience, face greater challenges in recovering from acute hemorrhage, thereby exacerbating their prognosis. Additionally, our multivariate logistic analysis identified TBV as an independent risk factor for adverse outcomes in EVT-treated aSAH patients, corroborating its centrality in the LightGBM model. The SHAP-derived cutoff values provided clear clinical insight, with TBV levels surpassing ~25mL serving as strong indicators of heightened poor outcome risks. This aligns with our prior research, which found TBV >20.4mL to be intimately linked with a significant surge in complication risks among aSAH patients.[3] The significance of TBV in poor-grade aSAH stems from its direct correlation with hemorrhage extent, which can lead to elevated intracranial pressure, severe cerebral vasospasm, and delayed cerebral ischemia.[31,32] These findings hold critical clinical implications, offering quantifiable thresholds for risk stratification and personalized treatment planning. Patients with TBV >25 mL may benefit from more aggressive perioperative management, such as early cerebrospinal fluid drainage to reduce intracranial hypertension, while those over 60 years old might benefit from enhanced multimodal supportive care and rehabilitation strategies to address age-related vulnerabilities. Incorporating these risk thresholds into clinical decision models could improve prognostic accuracy and aid in guiding individualized EVT strategies, ultimately improving patient outcomes.

Compared to traditional scoring systems, our model offers several distinct advantages. First, the employment of SHAP provides a comprehensive understanding of how each predictor influences the final outcome prediction. This transparency enhances clinical utility, particularly for junior clinicians, by enabling more precise identification of poor-grade aSAH patients and facilitating timely interventions. Moreover, SHAP analysis allows for individualized risk assessment by quantifying the impact of each predictor at the patient level, making it possible to provide personalized prognostic insights. This capability is particularly valuable in clinical settings, where tailored treatment strategies can significantly improve patient outcomes. With an AUC-ROC of 0.842 in the validation cohort, our model demonstrates strong predictive capability, surpassing or matching previous ML-based prognostic models for aSAH. Furthermore, the ability to continuously update the LightGBM model enhances its adaptability for clinical applications. By combining predictive power with explainability, our approach represents a significant step forward in bridging the gap between machine learning and clinical practice, advancing personalized prediction and precision medicine for poor-grade aSAH patients.

This study introduces a pioneering predictive ML model specifically designed for EVT-treated poor-grade aSAH patients—an area previously unexplored. As the first of its kind, this model accurately predicts post-EVT prognosis, enabling personalized management. However, several limitations must be acknowledged. Our data were derived from a registry cohort study, but patient enrollment was limited to a single center, necessitating multi-center validation for broader applicability. Second, while the Boruta algorithm effectively identified key predictive features, important variables beyond our dataset—such as genetic and molecular markers—may also influence prognosis. To enhance predictive accuracy, future research should integrate a broader range of biomarkers.

Conclusion

In poor-grade aSAH patients, endovascular coiling is an independent predictor of improved 12-month outcomes. The LightGBM model demonstrated strong predictive performance and generalizability across both training and validation cohorts. Utilizing SHAP algorithms enhanced the transparency and interpretability of the predictive models, facilitating clinical personalized decision-making. This study introduces a high-performance predictive model, providing clinicians with a valuable tool for accurately assessing prognosis in poor-grade aSAH patients undergoing endovascular treatment.

Abbreviations

EVT, Endovascular Treatment; aSAH, Aneurysmal Subarachnoid Hemorrhage; AUC-ROC, Area Under The Receiver Operating Characteristic Curve; SHAP, Shapley Additive Explanations; aOR, Adjusted Odds Ratios; ISAT, The International Subarachnoid Aneurysm Trial; ML, Machine Learning; CT, Computed Tomography; DSA, Digital Subtraction Angiography; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; EDC, Electronic Data Capture; WFNS, Neurosurgical Societies Scale; HH, the Hunt-Hess Grade; mFS, the Modified Fisher Scale; TBV, Total Bleeding Volume; IPH, Intraparenchymal Hemorrhage; IVH, Intraventricular Hemorrhage; CHD, Coronary Heart Disease; LR, Logistic Regression; DT, Decision Tree; Enet, Elastic Net; KNN, K-Nearest Neighbors; LightGBM, Light Gradient Boosting Machine; RF, Random Forest; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machines; MLP, Multilayer Perceptron; LASSO, Least Absolute Shrinkage and Selection Operator; IQR, Inter-quartile Range; DCA, Decision Curve Analysis; BS, Brier Score.

Data Sharing Statement

The corresponding author can provide the data supporting the findings of this study upon reasonable request.

Ethics Approval and Consent to Participate

The study protocol was approved by the main investigator institution, the Ethic Committee of The Second Affiliated Hospital of Nanchang University (IIT-O-2023-011), and all enrolled patients signed informed consent at admission. The present study complied with the Declaration of Helsinki.

Consent for Publication

All the authors consented to publication. This manuscript has not been published elsewhere and is not under consideration by another journal.

Funding

This research was supported by National Natural Science Foundation of China (Nos.81960456 and 82172989 for XGZ), Natural Science Foundation of Jiangxi Province Project (Nos. 20202BAB206053 for MJW, 20224BAB216074 and 20232BAB206085 for TFY), Training Program for academic and technical leaders in major disciplines in Jiangxi Province-Young Talents Project (No. 20225BCJ23024 for MJW).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Brian LH, Nerissa UK, Sepideh A-H, et al. 2023 guideline for the management of patients with aneurysmal subarachnoid hemorrhage: a guideline from the American Heart Association/American Stroke Association. *Stroke*. 2023;54(7):e314–70. doi:10.1161/str.0000000000000436
2. Molyneux AJ, Kerr RS, Yu LM, 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(9488):809–817. doi:10.1016/S0140-6736(05)67214-5
3. Hu P, Wu Y, Yan T, et al. Deep learning-based quantification of total bleeding volume and its association with complications, disability, and death in patients with aneurysmal subarachnoid hemorrhage. *J Neurosurg*. 2024;141(2):343–354. doi:10.3171/2024.1.JNS232280
4. Athanasios KP, Marcel KA, Jan FC, et al. Aneurysmal subarachnoid hemorrhage. *Dtsch Arztebl Int*. 2017;114(13):226. doi:10.3238/arztebl.2017.0226
5. Owen BS, Ofer S, Chen F, et al. Aneurysmal subarachnoid hemorrhage: trends, outcomes, and predictions from a 15-year perspective of a single neurocritical care unit. *Neurosurgery*. 2020;88(3):574–583. doi:10.1093/neuros/nyaa465
6. David AW, Peter N, Felipe CA, Cameron GM, Joseph MZ, Robert FS. Time course of recovery following poor-grade SAH: the incidence of delayed improvement and implications for SAH outcome study design. *J Neurosurg*. 2013;119(3):606–612. doi:10.3171/2013.4.Jns121287
7. Han L, Gaici X, Sisi L, et al. An accurate prognostic prediction for aneurysmal subarachnoid hemorrhage dedicated to patients after endovascular treatment. *Ther Adv Neurol Disord*. 2022;15. doi:10.1177/17562864221099473
8. Kaiwen W, Qingyuan L, Shaohua M, et al. A decision tree model to help treatment decision-making for severe spontaneous intracerebral hemorrhage. *Int J Surg*. 2023;110(2):788–798. doi:10.1097/js9.0000000000000852
9. Lei S, Hua Y, Yanze W, et al. Explainable machine learning in outcome prediction of high-grade aneurysmal subarachnoid hemorrhage. *Aging*. 2024;16(5):5618–5633. doi:10.18632/aging.205621
10. Fandi H, Qingqing Z, Wanwan Z, et al. A correlation and prediction study of the poor prognosis of high-grade aneurysmal subarachnoid hemorrhage from the neutrophil percentage to albumin ratio. *Clin Neurol Neurosurg*. 2023;230:107788. doi:10.1016/j.clineuro.2023.107788
11. Duo Y, George WW, David A, et al. Machine learning prediction of the adverse outcome for nontraumatic subarachnoid hemorrhage patients. *Ann Clin Transl Neurol*. 2020;7(11):2178–2185. doi:10.1002/acn3.51208
12. Scott ML, Gabriel E, Hugh C, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9
13. Hugh C, Scott ML, Su-In L. Explaining a series of models by propagating Shapley values. *Nat Commun*. 2022;13(1):4512. doi:10.1038/s41467-022-31384-3
14. Hu P, Zhou H, Yan T, et al. Deep learning-assisted identification and quantification of aneurysmal subarachnoid hemorrhage in non-contrast CT scans: development and external validation of Hybrid 2D/3D UNet. *NeuroImage*. 2023;279:120321. doi:10.1016/j.neuroimage.2023.120321
15. Hu P, Liu Y, Li Y, et al. A comparison of LASSO regression and tree-based models for delayed cerebral ischemia in elderly patients with subarachnoid hemorrhage. *Front Neurol*. 2022;13:791547. doi:10.3389/fneur.2022.791547
16. Eekhout I, de Boer R, Twisk J, de Vet H, Heymans M. Missing data: a systematic review of how they are reported and handled. *Epidemiology*. 2012;23(5):729–732. doi:10.1097/EDE.0b013e3182576cdb
17. Diringer MN, Bleck TP, Claude Hemphill JI, et al. Critical care management of patients following aneurysmal subarachnoid hemorrhage: recommendations from the neurocritical care society's multidisciplinary consensus conference. *Neurocrit Care*. 2011;15(2). doi:10.1007/s12028-011-9605-9
18. Jack H, Mohammed OD, Dhruv K, et al. Outcomes following poor-grade aneurysmal subarachnoid haemorrhage: a prospective observational study. *Acta Neurochir*. 2023;165(12):3651–3664. doi:10.1007/s00701-023-05884-0
19. Tatsuya I, Fusao I, Nao I, et al. Superiority of endovascular coiling over surgical clipping for clinical outcomes at discharge in patients with poor-grade subarachnoid hemorrhage: a registry study in Japan. *Neurosurgery*. 2023;94(5):1051–1060. doi:10.1227/neu.0000000000002782
20. Jinjin L, Ye X, Ming Z, et al. Predicting long-term outcomes after poor-grade aneurysmal subarachnoid hemorrhage using decision tree modeling. *Neurosurgery*. 2020;87(3):523–529. doi:10.1093/neuros/nyaa052
21. Shen J, Yu J, Huang S, et al. Scoring model to predict functional outcome in poor-grade aneurysmal subarachnoid hemorrhage. *Front Neurol*. 2021;12:601996. doi:10.3389/fneur.2021.601996
22. Khansa R, Adnan Q, Mohammed G, Ala A-F, Adeel R, Junaid Q. Explainable, trustworthy, and ethical machine learning for healthcare: a survey. *Comput Biol Med*. 2022;149:106043. doi:10.1016/j.compbiomed.2022.106043

23. Kipp WJ, Jessica TS, Benjamin SG, et al. Artificial Intelligence in Cardiology. *J Am Coll Cardiol*. 2018;71(23):2668–2679. doi:10.1016/j.jacc.2018.03.521
24. Jeremy P, Shuang D, Walter N. Opening the black box: the promise and limitations of explainable machine learning in cardiology. *Can J Cardiol*. 2021;38(2):204–213. doi:10.1016/j.cjca.2021.09.004
25. Khoa AT, Olga K, Andrew B, Elizabeth DW, John VP, Nicola W. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med*. 2021;13(1). doi:10.1186/s13073-021-00968-x
26. Mohanad MA, Freya A, Jung Won C, et al. Prediction of disease comorbidity using explainable artificial intelligence and machine learning techniques: a systematic review. *Int J Med Inform*. 2023;175:105088. doi:10.1016/j.ijmedinf.2023.105088
27. Jun Y, Yuetong X, Qian C, et al. LightGBM: accelerated genomically designed crop breeding through ensemble learning. *Genome Biol*. 2021;22(1):1–24. doi:10.1186/s13059-021-02492-y
28. Jiayi Z, Wenlong L, Huiquan Z, et al. Identifying dementia from cognitive footprints in hospital records among Chinese older adults: a machine-learning study. *Lancet Reg Health West Pac*. 2024;46. doi:10.1016/j.lanwpc.2024.101060
29. Carlina EV, Nicolaas AB, Jaqueline B, et al. Prediction of outcome after aneurysmal subarachnoid hemorrhage. *Stroke*. 2019;50(4):837–844. doi:10.1161/strokeaha.118.023902
30. Nicolai M, Victoria V, Isabel charlotte H, et al. External validation of the HATCH (Hemorrhage, Age, Treatment, Clinical State, Hydrocephalus) score for prediction of functional outcome after subarachnoid hemorrhage. *Neurosurgery*. 2022;91(6):906–912. doi:10.1227/neu.0000000000002128
31. Lagares A, Jiménez-Roldán L, Gomez P, et al. Prognostic value of the amount of bleeding after aneurysmal subarachnoid hemorrhage: a quantitative volumetric study. *Neurosurgery*. 2015;77(6):898–907. doi:10.1227/neu.0000000000000927
32. García S, Torné R, Hoyos J, et al. Quantitative versus qualitative blood amount assessment as a predictor for shunt-dependent hydrocephalus following aneurysmal subarachnoid hemorrhage. *J Neurosurg*. 2018;131(6):1743–1750. doi:10.3171/2018.7.Jns18816