Journal of Multidisciplinary Healthcare

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Construction of an Interpretable Model of the Risk of Post-Traumatic Brain Infarction Based on Machine Learning Algorithms: A Retrospective Study

Shaojie Li[1,*], Hongjian Li[2,*], Baofang Wu[1,*], Rujun Pan[3], Yuqi Liu[4], Jiayin Wang[1], De Wei[3], Hongzhi Gao[1]

[1]Department of Neurosurgery, the Second Affiliated Hospital of Fujian Medical University, Quanzhou, 362000, People's Republic of China; [2]School of Medical Imaging, North Sichuan Medical College, Nanchong, 634700, People's Republic of China; [3]Department of Neurosurgery, Fuzhou University Affiliated Provincial Hospital, Fuzhou, 350001, People's Republic of China; [4]Department of Intensive Care Unit, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, 362000, People's Republic of China

*These authors contributed equally to this work

Correspondence: De Wei; Hongzhi Gao, Email weidele@fjmu.edu.cn; gaohongzhi@fjmu.edu.cn

Background: Post-traumatic cerebral infarction (PTCI) is a severe complication resulting from traumatic brain injury (TBI), which can lead to permanent neurological damage or death. The investigation of the factors associated with PTCI and the establishment of predictive models are crucial for clinical practice.

Methods: We made a retrospective analysis of clinical data from 1484 TBI patients admitted to the Neurosurgery Department of a provincial hospital from January 2018 to December 2023. Predictive factors were identified using the Least Absolute Shrinkage and Selection Operator (LASSO) and multivariable logistic regression analysis. Several machine learning (ML) classification models were developed and compared. The interpretations of the ML models' predictions were provided by SHAP values.

Results: Key predictors included age, bilateral brain contusions, platelet count, uric acid, glucose, traumatic subarachnoid hemorrhage, and surgical treatment. The logistic regression (LR) model outperformed other ML algorithms, demonstrating superior performance in the test set with an AUC of 0.821, accuracy of 0.845, Matthews correlation coefficient (MCC) of 0.264, area under the receiver operating characteristic curve (AUROC) of 0.711, precision of 0.56, and specificity of 0.971. It had stable performance in the ten-fold cross-validation.

Conclusion: ML algorithms, integrating demographic and clinical factors, accurately predicted the risk of PTCI occurrence. Interpretations using the SHAP method offer guidance for personalized treatment of different patients, filling gaps between complex clinical data and actionable insights.

Keywords: traumatic brain injury, post-traumatic cerebral infarction, retrospective study, machine learning, prediction model

Introduction

Traumatic brain injury (TBI) includes impairments and pathological changes in brain function due to external forces, including concussions and traumatic herniations.[1] TBI has the highest incidence rate among common neurological disorders, bringing a significant global public health challenge with high morbidity and mortality rates.[2] Post-traumatic cerebral infarction (PTCI), a severe complication caused by TBI, can lead to permanent neurological damage or death.[3] Although PTCI occurs in only 1.9% to 20.3% of all TBI cases, it profoundly affects patient prognosis.[4] Brain injuries may cause direct vascular damage, changes in hemodynamics, or vascular compression due to increased intracranial pressure, triggering an infarction.[5] Clinically, the manifestations of PTCI depend on the affected cerebral vascular territory and may lead to motor and sensory impairments of different degrees. For instance, infarctions in the posterior cerebral artery (PCA) region are mainly associated with visual field defects and sensory disturbances.[6,7] As a "secondary injury" after head

Received: 25 October 2024
Accepted: 9 January 2025
Published: 16 January 2025

trauma, PTCI significantly increases patient disability and mortality rates. In TBI patients with severe consciousness disorders and other complications, typical symptoms and signs of infarction are often obscured, complicating early diagnosis.[8] Therefore, timely identification of these risk factors and the development of effective predictive models are crucial for preventing PTCI and implementing early preventive measures.

In recent years, machine learning (ML) technologies have significantly advanced in the healthcare sector, particularly demonstrating substantial potential in disease prediction, diagnosis, and treatment decision support.[9,10] By processing and analyzing large-scale clinical data, ML can identify complex patterns and predict disease progression.[11] This capability is particularly crucial for understanding and predicting the risk of PTCI, which is closely associated with various risk factors and biomarkers. In clinical practice, the early diagnosis of PTCI often relies on limited imaging data and patient condition observations and lacks comprehensive analysis of extensive historical data.[12] Consequently, ML-based models provide a more accurate and objective method for assessing PTCI risk. Moreover, by introducing interpretative tools such as Shapley Additive exPlanations (SHAP), ML models can not only deliver predictive outcomes but also identify the key drivers affecting predictions, thus offering clearer guidance for clinical decision-making.

This study aims to develop and validate a machine-learning-based PTCI risk prediction model that improves prediction accuracy by integrating patient data in multiple dimensions and using precise algorithmic analyses. We anticipate that this work will provide a scientific basis for clinical decision-making, so that physicians can identify high-risk PTCI patients earlier and manage them individually, thereby reducing the incidence rate of PTCI and improving the quality of life of patients.

Materials and Methods

Materials

This article describes an observational and retrospective cohort study involving patients with traumatic brain injuries treated in the Neurosurgery Department of the Affiliated Provincial Hospital of Fuzhou University from January 2018 to December 2023. The study was approved by the Ethics Committee of the Provincial Hospital of Fuzhou University (K2024-06-057). Given the retrospective and observational nature of the study, informed consent was waived.

The inclusion criteria are as follows: 1) patients admitted within 24 hours after injury; 2) patients admitted for the first time to our hospital without prior treatment at other hospitals; 3) initial head CT performed within 6 hours of admission showing intracranial hemorrhage but no stroke; 4) the main diagnosis of brain injury, with any other injuries having an Acute Injury Score (AIS) of less than 3 and no open wound. The exclusion criteria include: 1) patients under the age of 18; 2) patients with a history of atrial fibrillation, stroke, venous thrombosis, liver disease, or hematologic conditions; 3) history of anticoagulant use; 4) patients who experienced shock during treatment; 5) patients who died before a second head CT could be performed; 6) patients with incomplete data.

Data Collection

In this study, clinical and laboratory data on TBI patients was collected by reviewing relevant literature. Data included demographic variables (such as age, gender, smoking, drinking, diabetes, and history of hypertension) and baseline characteristics at admission (such as systolic and diastolic pressure, unequal pupil diameter, and Glasgow Coma Scale score). Laboratory indicators encompassed white blood cells, neutrophils, hemoglobin, platelets, monocytes, lymphocytes, high-density and low-density lipoproteins, urea nitrogen, creatinine, uric acid, albumin, cholesterol, triglycerides, fasting blood glucose, International Normalized Ratio (INR), Activated Partial Thromboplastin Time (APTT), Prothrombin Time (PT), D-dimer (DD), Systemic Inflammatory Index (SII), and Systemic Inflammatory Response Index (SIRI). Treatment measures included medications to reduce intracranial pressure and surgical interventions. Through CT imaging, we recorded the type and location of anatomical injuries, including lobar contusions, epidural and subdural hematomas, traumatic subarachnoid hemorrhage, intraventricular hemorrhage, fractures of the skull base and calvaria, multiple skull fractures, brain herniation, and cerebral infarction.

Diagnosis of PTCI

The diagnosis of PTCI was made through collaboration between neuroradiology and neurosurgery. Any well-defined area of low attenuation matching the distribution of arterial vasculature that appeared on any CT scan of the brain within two weeks of the accident was defined as a PTCI (Figure 1A-C).[4,13,14] The diagnosis of cerebral infarction was modified if subsequent studies showed that the diagnosis of cerebral infarction was related to other causes of evolving contusions, artifacts, etc. Then, the diagnosis of cerebral infarction was revised.

Statistical Analysis

For continuous variables that are normally distributed, we use independent samples *T*-tests for group comparisons, whereas non-parametric tests are used for non-normally distributed variables. Chi-square tests are done for categorical variables. To address collinearity and eliminate potential confounders, the variable is initially selected using LASSO regression and multivariable logistic regression. The dataset is then randomly divided into a training set and a test set in a ratio of 7:3, where the training set is used for building models and the test set is used for evaluating them. Models including XGBoost, Logistic regression, RandomForest, AdaBoost, and Naive Bayes are developed and assessed using AUC, calibration curves, and decision curve analysis (DCA). The best-performing model is selected. Furthermore, to verify the stability and reliability of the chosen model, 10-fold cross-validation is performed, comparing metrics such as accuracy, prevalence, recall, F1-Score, Matthews correlation coefficient, the area under the ROC curve (AUROC), precision, and specificity. Lastly, SHAP values are used to elucidate the importance of each feature and their specific contributions to the prediction outcomes, enhancing the interpretability of the model results.

Results

During the study, 2178 patients diagnosed with TBI were initially included. After excluding 694 patients for various reasons (Figure 2), data from 1484 patients were analyzed. Overall, the mean age of traumatic brain injury was 55.36 ± 17.33. Of these, 66.44% were male. The dataset was randomly divided into two parts: 70% (n=1038) for model training and 30% (n=446) for model validation. Between the training and validation sets, there was no statistically significant difference ($P > 0.05$). The incidence rate of PTCI was 15.90% in the training set and 16.14% in the validation set, with an overall incidence rate of 15.97% in the entire dataset. Detailed characteristics of the participants are shown in Table 1.

Screening of Factors Characterizing the Risk of Post-Traumatic Cerebral Infarction

LASSO regression analysis was made on all independent variables. PTCI was set as the dependent variable, in order to determine the risk factors connected to PTCI (Figure 3A and B). LASSO reduces overfitting by compressing variable coefficients and effectively addresses multicollinearity issues. The results showed Lambda = −4.756 at the minimum

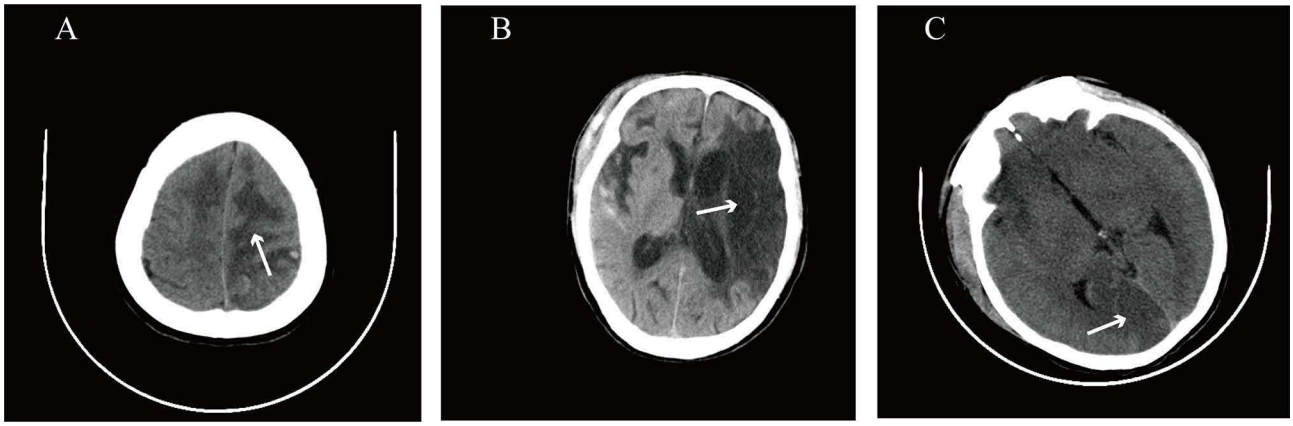

Figure 1 Different locations of cerebral infarction in CT after traumatic brain injury. The area indicated by the white arrow relative to the surrounding black area represents the area of cerebral infarction. (**A**). Low-density infarct lesions in the region innervated by the anterior cerebral artery; (**B**). Low-density infarct lesions in the region innervated by the middle cerebral artery; (**C**). Low-density infarct lesions in the region innervated by the posterior cerebral artery.

https://doi.org/10.2147/JMDH.S498420

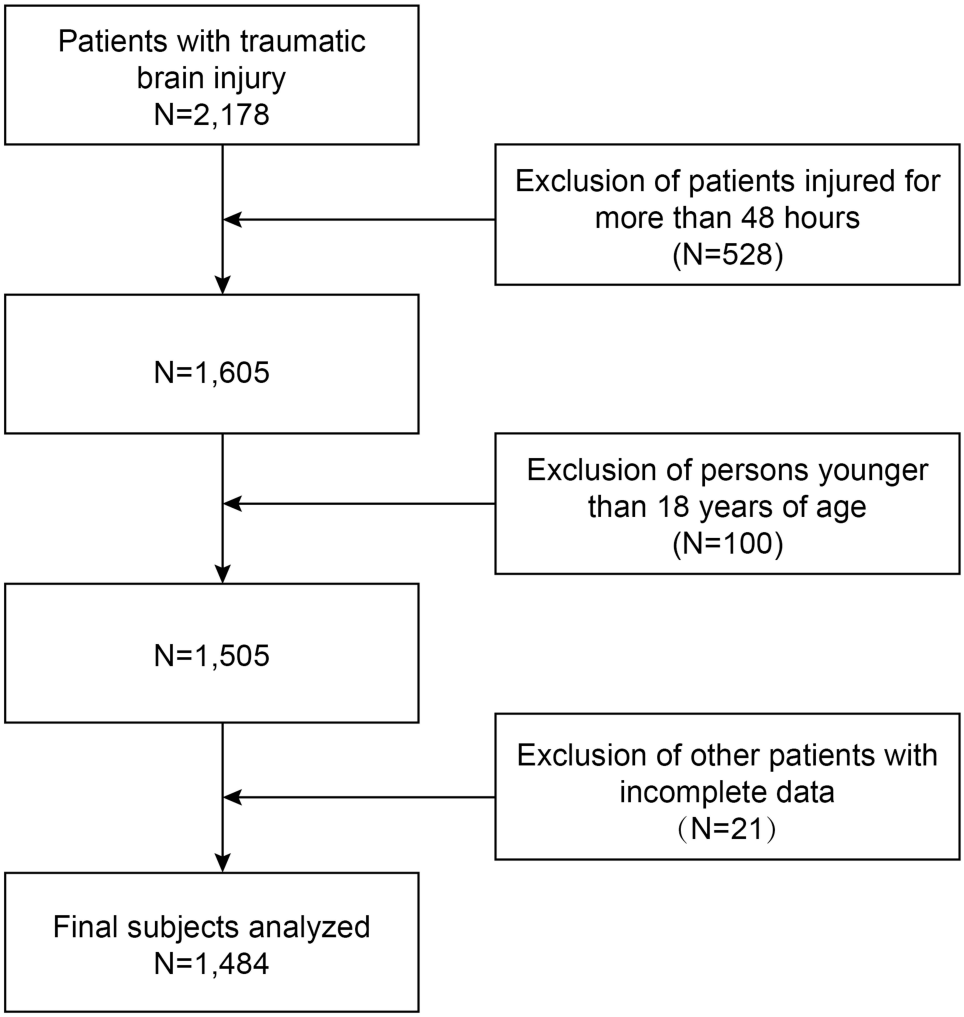

Figure 2 Participant Selection Flowchart.

mean square error, at which point the 48 independent variables were filtered down to 16, and these included age, history of hypertension, history of diabetes, intraventricular hemorrhage, traumatic subarachnoid hemorrhage, epidural hematoma, bilateral cerebral contusion, calvarial fracture, multiple skull fractures, surgery, neutrophils, platelets, creatinine, uric acid, albumin, and glucose. To further eliminate the impact of confounding factors, multivariable logistic regression was made on the selected variables. Ultimately, age, uric acid, glucose, surgery, bilateral cerebral contusion, platelets, and traumatic subarachnoid hemorrhage were identified as independent risk factors ($p < 0.05$), as shown in Figure 4.

Categorical Multi-Model Synthesis Analysis

In the training set, we trained the XGBoost, LR, RandomForest, AdaBoost, and Naive Bayes (NB) models. Each of them was repeated 10 times to establish the models. To evaluate the predictive performance, we used the AUC as the assessment metric. The results indicated that all models performed well in the training set (AUC > 0.7). However, in the test set, except for the Logistic Regression model which maintained its stability (training set AUC:0.833; test set AUC:0.821), the accuracy of other models declined, showing signs of overfitting (Figure 5A and B). Moreover, to assess the calibration and clinical applicability of the models, calibration curves and DCA were drawn. Compared to others, the Logistic Regression model's calibration curve was closer to the perfect 45-degree line and achieved a lower Brier score, indicating better consistency (Figure 5C). Additionally, this model

Table 1 Baseline Features in the Training and Test Sets

Variable Names	Train (N=1038)			Test (N=446)			Total
	Non-PTCI (N=873)	PTCI (N=165)	*P*-value	Non-PTCI (N=374)	PTCI (N=72)	*P*-value	*P*-value
GCS	13.205±3.334	13.812±2.441	0.026	13.187±3.376	13.569±3.039	0.372	0.775
Age (year)	52.464±16.765	71.03±10.957	<0.001	52.532±16.836	69.264±11.886	<0.001	0.853
Systolic blood pressure (mmHg)	135.545±21.323	144.697 ±20.712	<0.001	136.409±19.656	139.028 ±19.734	0.301	0.887
Diastolic blood pressure (mmHg)	78.691±11.81	80.297±12.028	0.11	79.594±13.192	76.472±9.054	0.055	0.834
SII	1969.706 ±1827.458	1717.689 ±1913.635	0.107	2004.845 ±1655.785	1505.533 ±1424.992	0.017	0.957
SIRI	6.765±8.282	5.113±4.826	0.013	7.749±11.183	4.376±4.054	0.012	0.155
Neutrophil (10^9/L)	9.856±4.526	8.256±3.58	<0.001	10.25±5.083	7.403±3.084	<0.001	0.467
Blood platelet (10^9/L)	196.027±61.615	190.827 ±63.978	0.323	195.602±60.267	196.528 ±59.802	0.905	0.874
Hemoglobin (g/L)	127.418±20.39	121.937 ±19.217	0.001	127.228±20.451	120.722 ±17.856	0.012	0.748
Monocyte (10^9/L)	0.631±0.302	0.57±0.236	0.014	0.667±0.364	0.559±0.221	0.016	0.109
Lymphocyte (10^9/L)	1.326±0.655	1.273±0.658	0.336	1.28±0.621	1.256±0.654	0.768	0.254
Leucocyte (10^9/L)	11.886±4.563	10.182±3.6	<0.001	12.294±5.218	9.279±3.065	<0.001	0.465
HDL-C	1.253±0.337	1.27±0.334	0.553	1.272±0.354	1.264±0.316	0.87	0.43
LDL-C	2.558±0.744	2.625±0.914	0.31	2.629±0.77	2.55±0.776	0.427	0.276
Urea nitrogen (mmol/l)	5.075±1.96	5.859±2.979	<0.001	4.948±2.163	5.372±2.003	0.124	0.135
Creatinine (umol/l)	71.966±29.17	80.765±89.886	0.021	70.185±28.288	72.146±24.323	0.582	0.211
Uric acid (umol/l)	310.284±110.13	294.773 ±95.988	0.091	312.012±110.595	302.567 ±106.143	0.505	0.664
Albumin (mg/L)	40.251±4.954	38.671±4.816	<0.001	39.912±5.344	38.546±3.577	0.038	0.278
Glucose (mmol/l)	7.005±2.745	7.064±2.624	0.8	7.252±2.86	6.559±1.96	0.05	0.417
Triglyceride (mmol/l)	4.16±0.841	4.206±1.014	0.53	4.254±0.888	4.14±0.902	0.322	0.167
Triglyceride (mmol/l)	1.321±1.557	1.287±1.467	0.792	1.37±1.814	1.057±0.543	0.148	0.965
INR	1.036±0.157	1.058±0.174	0.098	1.052±0.235	1.022±0.121	0.291	0.436
APTT	29.128±5.398	29.775±5.614	0.161	29.69±7.885	28.689±3.879	0.293	0.388
PT	12.136±1.768	12.361±1.929	0.14	12.292±2.443	11.936±1.42	0.232	0.572
D-dimer	7.659±10.033	7.155±9.26	0.549	7.577±9.285	5.487±8.176	0.076	0.537
Gender (%)			0.197				
No	280 (32.07)	62 (37.58)		125 (33.42)	31 (43.06)	0.151	0.484
Yes	593 (67.93)	103 (62.42)		249 (66.58)	41 (56.94)		
Smoking history (%)			0.415			0.483	0.434
No	783 (89.69)	152 (92.12)		329 (87.97)	66 (91.67)		
Yes	90 (10.31)	13 (7.88)		45 (12.03)	6 (8.33)		
Drinking history (%)			0.452			1	0.517
No	802 (91.87)	155 (93.94)		340 (90.91)	66 (91.67)		
Yes	71 (8.13)	10 (6.06)		34 (9.09)	6 (8.33)		
History of hypertension (%)			<0.001			<0.001	0.589
No	684 (78.35)	88 (53.33)		286 (76.47)	39 (54.17)		
Yes	189 (21.65)	77 (46.67)		88 (23.53)	33 (45.83)		
History of diabetes (%)			<0.001			0.024	0.669
No	763 (87.40)	125 (75.76)		323 (86.36)	54 (75.00)		
Yes	110 (12.60)	40 (24.24)		51 (13.64)	18 (25.00)		

(*Continued*)

https://doi.org/10.2147/JMDH.S498420

Table 1 (Continued).

Variable Names	Train (N=1038)			Test (N=446)			Total
	Non-PTCI (N=873)	PTCI (N=165)	*P*-value	Non-PTCI (N=374)	PTCI (N=72)	*P*-value	*P*-value
Unequal pupil diameters (%)			0.412			1	0.841
No	410 (47.07)	84 (50.91)		174 (46.90)	34 (47.22)		
Yes	461 (52.93)	81 (49.09)		197 (53.10)	38 (52.78)		
Frontal lobe injury (%) (%)			1			0.037	0.659
No	516 (59.11)	98 (59.39)		218 (58.29)	52 (72.22)		
Yes	357 (40.89)	67 (40.61)		156 (41.71)	20 (27.78)		
Temporal lobe injury (%)			0.127			0.243	0.266
No	601 (68.84)	124 (75.15)		268 (71.66)	57 (79.17)		
Yes	272 (31.16)	41 (24.85)		106 (28.34)	15 (20.83)		
Parietal lobe injury (%)			0.683			0.492	0.378
No	815 (93.36)	152 (92.12)		341 (91.18)	68 (94.44)		
Yes	58 (6.64)	13 (7.88)		33 (8.82)	4 (5.56)		
Occipital lobe injury (%)			0.519			0.819	1
No	852 (97.59)	159 (96.36)		364 (97.33)	71 (98.61)		
Yes	21 (2.41)	6 (3.64)		10 (2.67)	1 (1.39)		
Remaining injuries (%)			1			0.921	0.114
No	838 (95.99)	158 (95.76)		365 (97.59)	71 (98.61)		
Yes	35 (4.01)	7 (4.24)		9 (2.41)	1 (1.39)		
Ventricular hemorrhage (%)			0.33			0.09	0.275
No	851 (97.48)	158 (95.76)		362 (96.79)	66 (91.67)		
Yes	22 (2.52)	7 (4.24)		12 (3.21)	6 (8.33)		
Subarachnoid hemorrhage (%)			0.253			0.001	0.82
No	199 (22.79)	45 (27.27)		79 (21.12)	29 (40.28)		
Yes	674 (77.21)	120 (72.73)		295 (78.88)	43 (59.72)		
Epidural hematoma (%)			0.002			0.019	0.817
No	684 (78.35)	147 (89.09)		289 (77.27)	65 (90.28)		
Yes	189 (21.65)	18 (10.91)		85 (22.73)	7 (9.72)		
Skull base fracture (%)			0.025			0.378	0.164
No	807 (92.44)	161 (97.58)		338 (90.37)	68 (94.44)		
Yes	66 (7.56)	4 (2.42)		36 (9.63)	4 (5.56)		
Skull cap fracture (%)			0.001			0.002	0.177
No	555 (63.57)	127 (76.97)		219 (58.56)	57 (79.17)		
Yes	318 (36.43)	38 (23.03)		155 (41.44)	15 (20.83)		
Subdural hemorrhage (%)			0.183			0.575	0.946
No	444 (50.86)	74 (44.85)		188 (50.27)	33 (45.83)		
Yes	429 (49.14)	91 (55.15)		186 (49.73)	39 (54.17)		
Bilateral cerebral contusions (%)			0.036			0.53	0.985
No	414 (47.42)	63 (38.18)		174 (46.52)	30 (41.67)		
Yes	459 (52.58)	102 (61.82)		200 (53.48)	42 (58.33)		
Cerebral hernia (%)			0.157			0.39	0.67
No	826 (94.62)	161 (97.58)		351 (93.85)	70 (97.22)		
Yes	47 (5.38)	4 (2.42)		23 (6.15)	2 (2.78)		
Multiple skull fractures (%)			0.004			0.037	0.73
No	797 (91.29)	162 (98.18)		338 (90.37)	71 (98.61)		
Yes	76 (8.71)	3 (1.82)		36 (9.63)	1 (1.39)		

(*Continued*)

Table 1 (Continued).

Variable Names	Train (N=1038)			Test (N=446)			Total
	Non-PTCI (N=873)	PTCI (N=165)	*P*-value	Non-PTCI (N=374)	PTCI (N=72)	*P*-value	*P*-value
Surgeries (%)			0.005			0.004	0.554
No	712 (81.56)	150 (90.91)		296 (79.14)	68 (94.44)		
Yes	161 (18.44)	15 (9.09)		78 (20.86)	4 (5.56)		
Dehydration drug use (%)			0.345			0.614	0.269
No	502 (57.50)	102 (61.82)		203 (54.28)	42 (58.33)		
Yes	371 (42.50)	63 (38.18)		171 (45.72)	30 (41.67)		

Note: Continuous variables are described using means and standard deviations, while categorical variables are represented using frequencies and percentages.
Abbreviations: PTCI, Post-traumatic cerebral infarction; GCS, Glasgow Coma Scale; SII, Systemic Immune Inflammation Index; SIRI, Systemic Immune Inflammation Index; HDL-C, High density lipoprotein cholesterol; HDL-C, Low-Density Lipoprotein Cholesterol; INR, international normalized ratio; APTT, Activated partial thromboplastin time; PT, prothrombin time;

had a higher net benefit in the decision curve analysis, covering the broadest threshold probability range from 0% to 73% (Figure 5D). Based on these findings, the LR model can be considered as the optimal model.

Evaluation of the Best Models

To improve the reliability of model evaluation, in this study, 10-fold cross-validation was made on the LR model (Figure 6A and B). The results indicated that the average AUCs for the training and test sets were 0.816 (0.786–0.847) and 0.804 (0.750–0.859) respectively, demonstrating that the model had a stable and accurate predictive performance on both datasets (Figure 6C and D). Additionally, in the study, various performance metrics were calculated and compared, including Accuracy, Recall, F1-Score, Matthews correlation coefficient (MCC), AUROC, Precision, and Specificity. In the test set, the LR model surpassed other models in terms of Accuracy, Prevalence, MCC, AUROC, Precision, and Specificity, showing outstanding performance (Figure 7) (Supplementary Table 1).

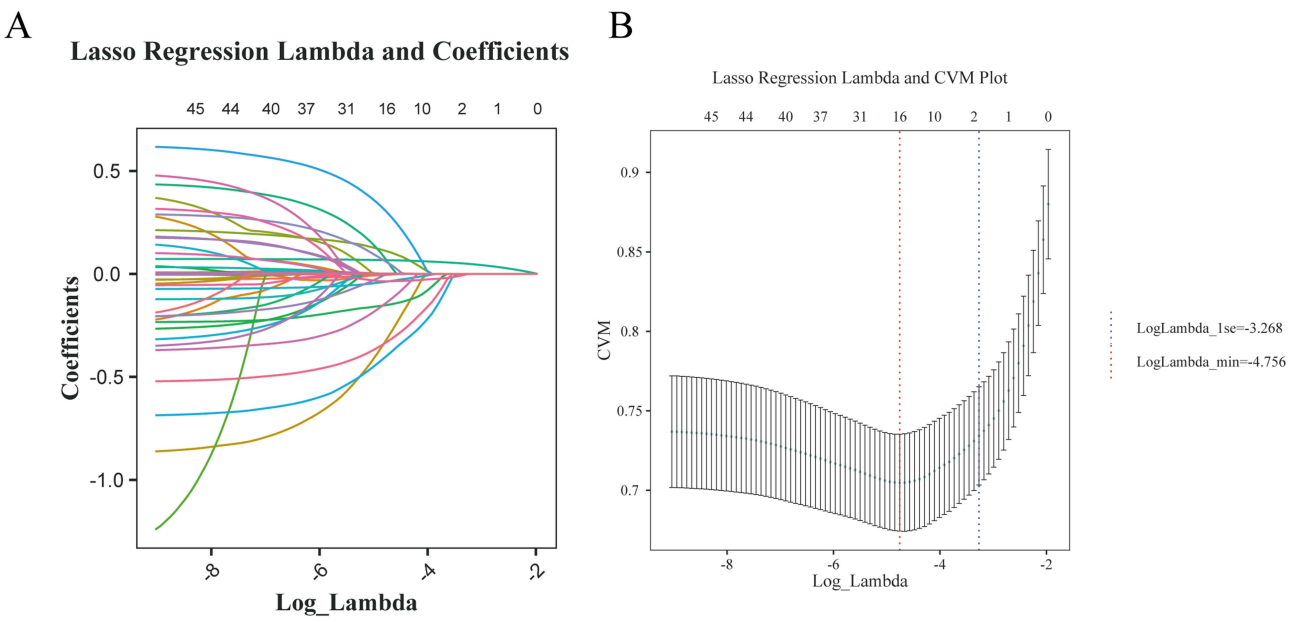

Figure 3 Selection of characteristic factors using LASSO regression analysis. (**A**) A relationship curve showing the link between log (lambda) and partial likelihood deviation (binomial deviation). (**B**) The 48 features' LASSO coefficient profiles. Using the minimal criteria and the one standard error (SE) of the minimum criteria (the 1-SE criteria), dot vertical lines were created at the optimal levels.

https://doi.org/10.2147/JMDH.S498420

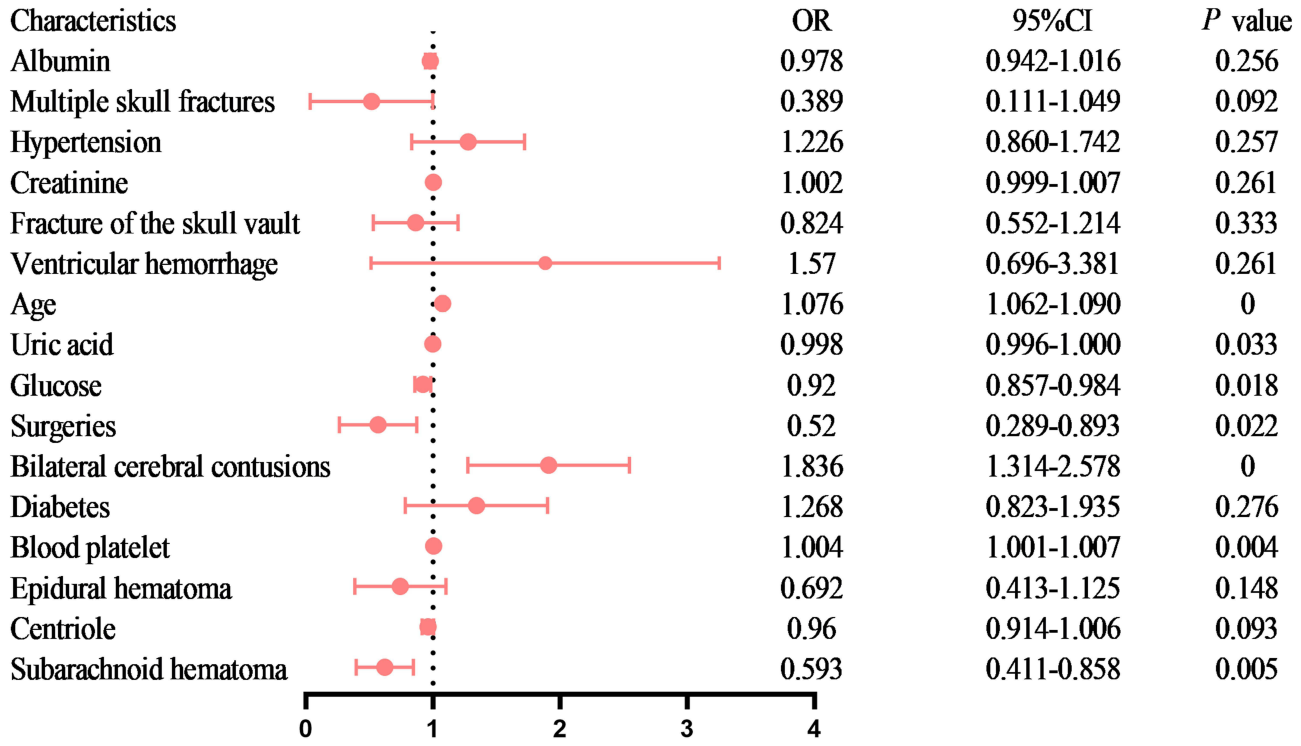

Figure 4 Multifactor regression results for variables.

Model Interpretation

To intuitively explain the impact of selected variables on the model, in this study, SHAP (Shapley Additive exPlanations) values were utilized to elucidate their contributions to the PTCI prediction model. As illustrated in Figure 8A, the horizontal position indicates whether the SHAP value increases or decreases the probability of prediction. Red indicates high values and blue indicates low values. Results indicated that increased age, bilateral brain contusions, and elevated platelet counts positively contribute to PTCI prediction, whereas uric acid, glucose, traumatic subarachnoid hemorrhage, and surgical treatment have a negative impact. Variables are ranked on the vertical axis by importance. Additionally, the study presents individual force plots for a non-PTCI (Figure 8B) and a PTCI patient (Figure 8C) to demonstrate the model's interpretability. SHAP values reveal the key features affecting individual predictions and their specific contributions to PTCI prediction, where red features increase the PTCI risk and blue features decrease the PTCI risk. The length of the arrows helps to visualize the magnitude of impact on the prediction; the longer the arrow is, the greater the effect is.

Discussion

In this study, the incidence rate of PTCI in patients with TBI was 15.97%, which was consistent with the previously reported rates of 1.9% to 20.3%. The variance in the incidence rate of PTCI may be attributed to factors such as medical resources, population health baselines, and trauma management and care approaches. This study is the first to apply machine learning techniques to predict the risk of PTCI. Feature selection was refined through LASSO regression and multivariable logistic regression analysis, effectively eliminating collinearity and potential confounders among variables. Age, uric acid levels, blood glucose, surgical treatment, bilateral brain contusions, platelet count, and traumatic subarachnoid hemorrhage were identified as independent risk factors for PTCI. Based on these findings, five machine-learning models were developed and validated to predict PTCI risk. Among them, the logistic regression model had the highest AUC, good calibration, and net benefit. Additionally, this model had superior performance in various evaluations, including ten-fold cross-validation. The results indicate that the logistic regression model performs well and generalizes

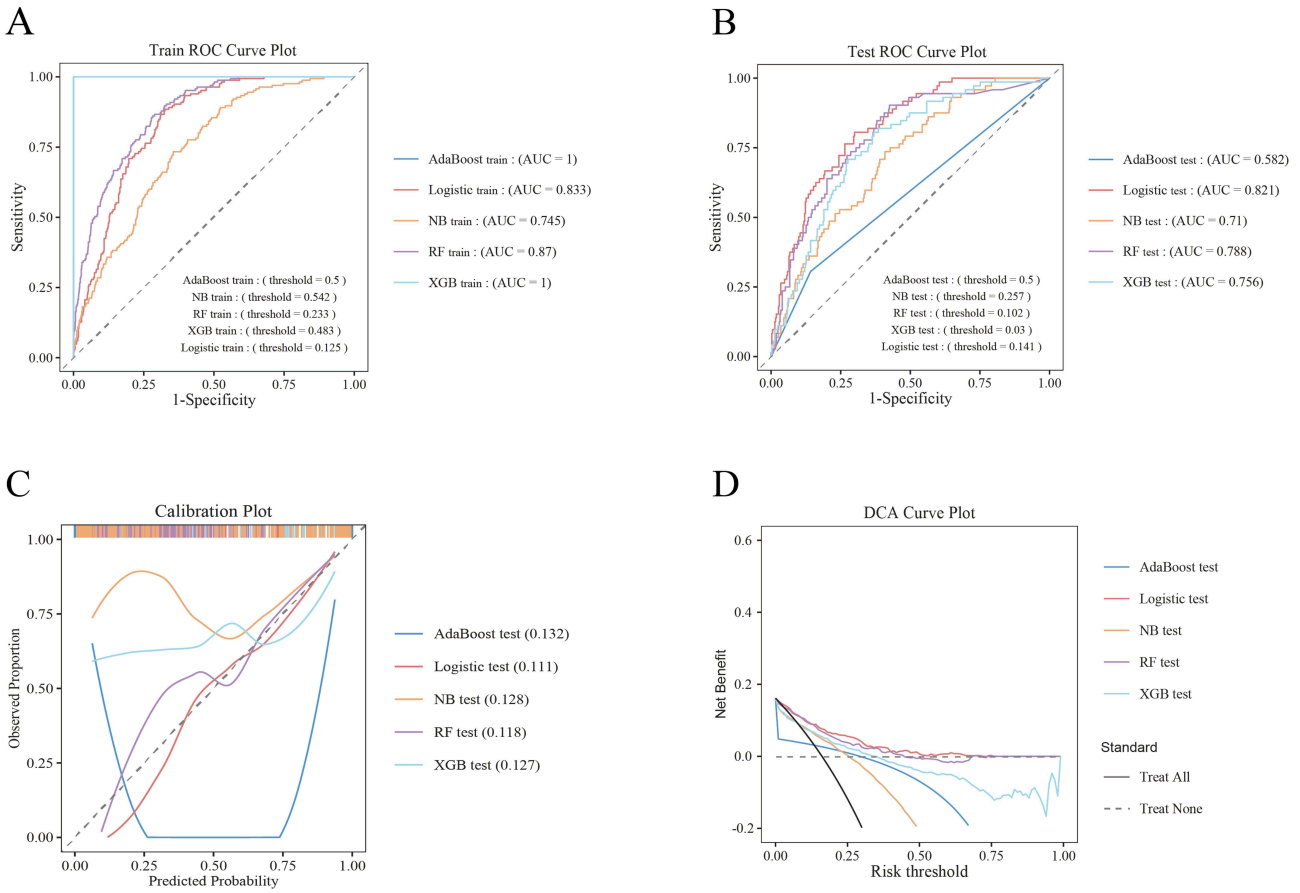

Figure 5 Comprehensive performance of several machine learning. (**A**) ROC curves for 5 models in the train set. (**B**) ROC curves in the test set. (**C**) Calibration of the 5 models in the test set. (**D**) DCA curves for 5 models in the test set.

effectively in the prediction of PTCI. Therefore, machine learning offers a promising new tool for handling complex clinical data, helping to fill existing knowledge and technology gaps.

In previous studies, the risk factors associated with PTCI have been explored using various statistical methods and target populations. In a retrospective study involving 353 patients with traumatic brain injuries, logistic regression analysis identified GCS, systolic pressure, brain herniation, and surgery as independent risk factors for PTCI after adjusting for confounders. Additionally, baseline data analysis indicated that most TBI patients could experience cerebral infarction within two weeks after injury.[15] Another study restricted its population to patients undergoing surgery for traumatic brain herniation, using classification and regression tree models to analyze predictive variables. The results suggested that the shock index, aspiration pneumonia, surgery, SAH, and intracranial pressure monitoring could be used to predict outcomes, with decreasing importance. Hu et al suggested that decompressive craniectomy could reverse brain herniation by alleviating malignant intracranial pressure, thereby preventing the progression of pulmonary inflammation and reducing the mortality rate of PTCI patients, aligning with our findings.[16] Wu et al included decompressive craniectomy, traumatic subarachnoid hemorrhage, hypotensive shock, and risk factors such as admission GCS, skull base fractures, and brain herniation in their study. However, they only made ROC analysis on individual variables. The overall predictive efficacy of the model was unclear. Nevertheless, no significant correlation between age and PTCI risk was observed.[17] Our study suggests that these differences may be related to regional treatment levels and sample sizes.

In this study, we used SHAP values to elucidate the outcomes of our machine learning model. By using this approach, we could specify the distinct contributions of individual predictors to model decisions, thereby enhancing the model's transparency and interpretability. We identified independent risk factors for PTCI, including age, uric acid levels, blood glucose, surgical interventions, bilateral brain contusions, platelet count, and traumatic subarachnoid hemorrhage. The

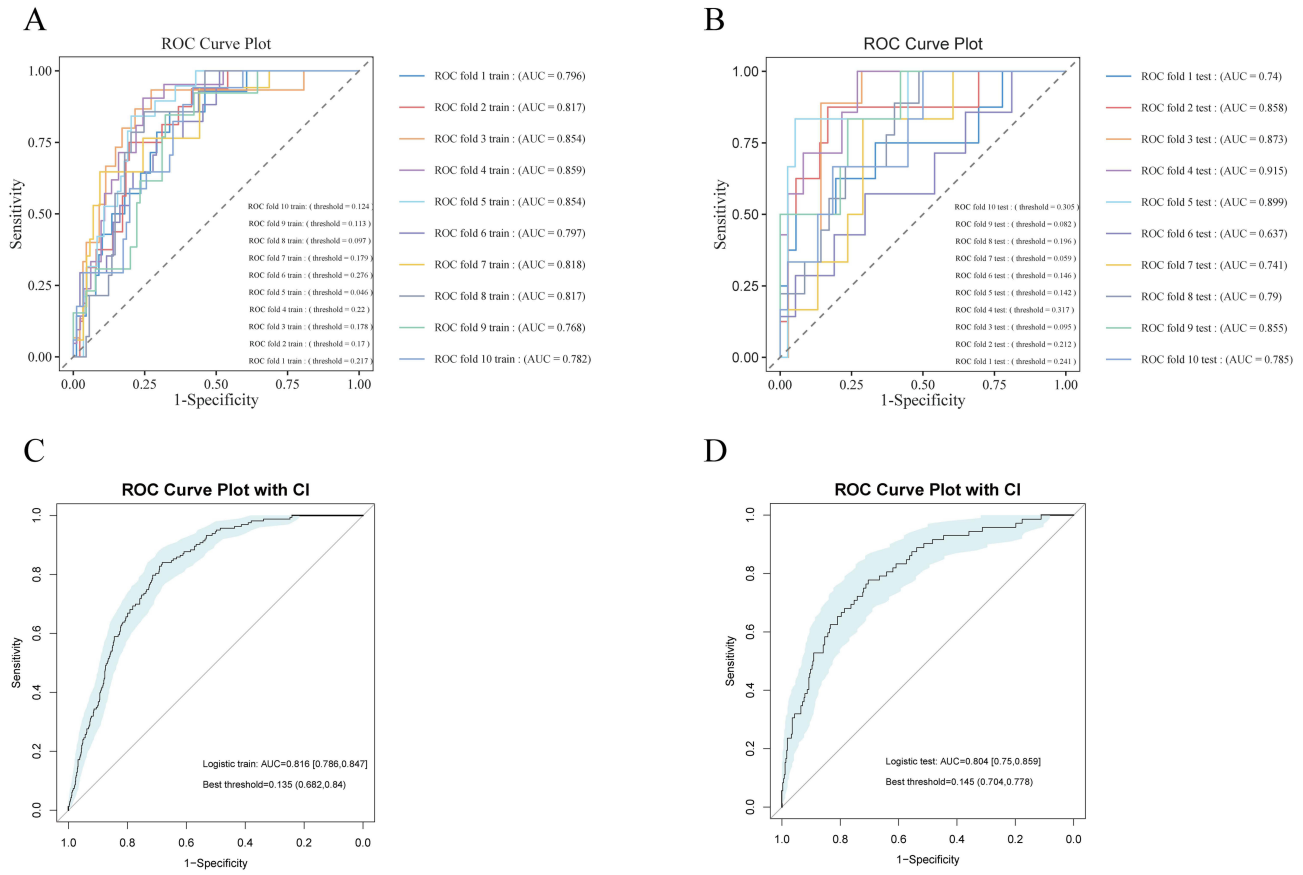

Figure 6 Ten-fold cross-validation results of the optimal model in the training and test sets. (**A**) ROC curves in the train set (**B**) ROC curves in the test set. (**C**) Average ROC of LR models in the train set. (**D**) Average ROC of LR models in the test set.

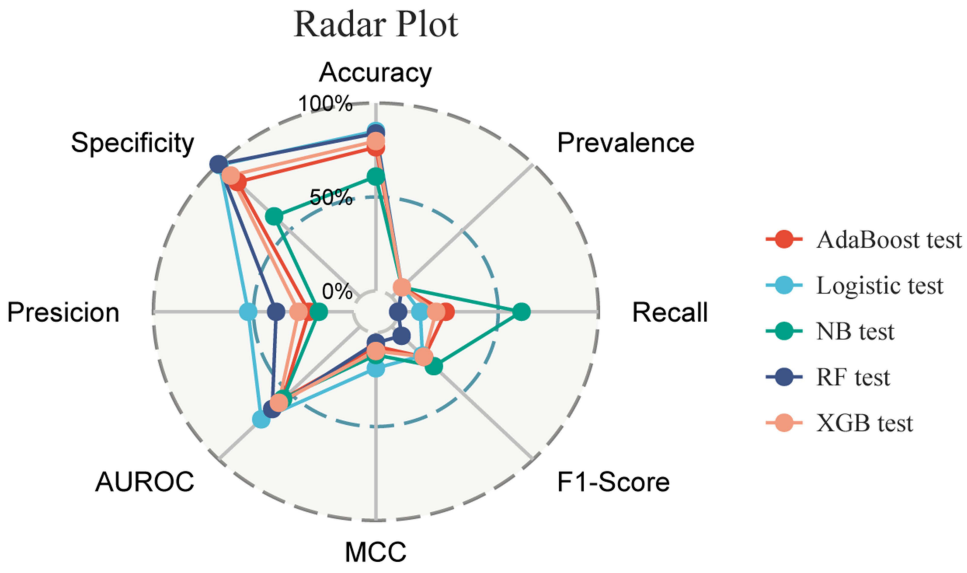

Figure 7 Comparison of several machine learning in terms of performance metrics.

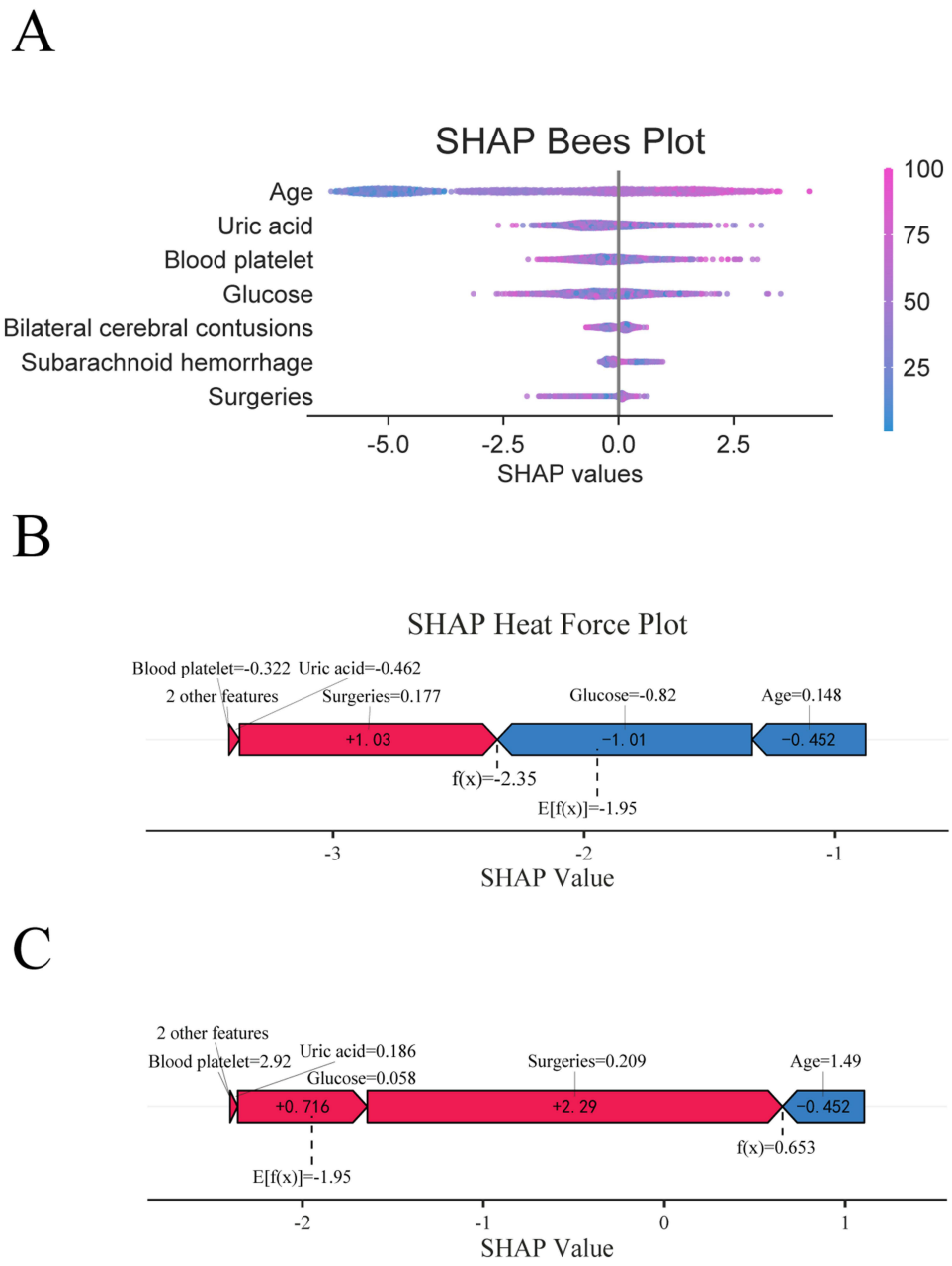

Figure 8 Interpretation of the model using SHAP values. (**A**) SHAP values for several variables. (**B**) Individual force plot for non-PTCI patients. (**C**) Individual force plot for PTCI patients.

presumed mechanisms are as follows: With the global increase in the elderly population, advanced age becomes a major risk factor for atherosclerotic cardiovascular diseases. Vascular pathologies such as arteriosclerosis, thinning of vessel walls, and decreased elasticity, are direct consequences of aging.[18] These changes may enhance vascular reactivity to trauma, increasing susceptibility to vascular damage and subsequent thrombosis, thereby increasing the risk of cerebral infarction.[4,19,20] A randomized controlled trial on uric acid highlighted its neuroprotective role in acute ischemic stroke patients through its potent antioxidant activity. During acute ischemic strokes (AIS), ischemia in brain tissues leads to the production of a substantial number of free radicals, such as hydroxyl radicals (·OH), hydrogen peroxide (H_2O_2), and peroxynitrite ($ONOO^-$), which further damage brain cells.[21] Uric acid can scavenge these free radicals, reducing oxidative stress and thus preventing further damage to brain cells.[22] Additionally, in patients with efficient collateral circulation, uric acid can reach the damaged brain areas through these routes, reducing cell death and thus playing

https://doi.org/10.2147/JMDH.S498420

a preventative role.[23] Moreover, traumatic brain injury activates the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system, leading to the production of large amounts of stress hormones such as catecholamines, cortisol, and glucagon. These hormones enhance glycogenolysis and gluconeogenesis, leading to hyperglycemia.[24] In a hyperglycemic state, pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6 are released in large quantities, exacerbating the inflammatory response.[25] This inflammation not only intensifies cerebral edema and blood-brain barrier disruption but also causes local cerebral ischemia through microvascular damage and thrombosis, increasing the risk of cerebral infarction.[26] Su et al suggested that bilateral brain contusions, often accompanied by extensive cerebral edema, significantly increased intracranial pressure and potential midline shifts, compressing cerebral vessels—particularly the middle cerebral and posterior cerebral arteries—reducing local blood flow and causing secondary cerebral infarctions.[27] Following head trauma, endothelial cells in vessels are damaged, and platelets gather around the injured vessels, activating a cascade of coagulation factors and increasing the risk of thrombosis.[28] Additionally, the gathering of platelets increases blood viscosity, raising the risk of microvascular and small-to-medium artery blood flow stasis, which can induce local ischemia in patients with traumatic brain injuries and poor cerebral perfusion more easily.[29] Cerebral vasospasm is one of the most common complications of traumatic subarachnoid hemorrhage (SAH), leading to severe ischemic damage.[30,31] In some guidelines, the use of vasodilators, such as nimodipine, after traumatic subarachnoid hemorrhage is recommended to reduce the incidence rate of vasospasm.[32,33] Therefore, in aggressive treatment protocols, SAH can serve as a trigger for active and preventive interventions, playing a protective role in preventing cerebral infarction.

This study has the following advantages. First, this is the first time that machine learning techniques are used to predict PTCI and explain the ML model through SHAP values. Second, the robustness and accuracy of the ML model are ensured by using a sufficiently large sample size and training the model multiple times by partitioning the training set and test set and using cross-validation methods. Despite the results of this study, there are limitations. First, the study included patients with mild traumatic brain injury to construct a predictive model applicable to a broader population. However, by introducing patients with mild traumatic craniocerebral injuries, bias may be introduced and the results need to be interpreted with caution. Second, because this study is a single-center retrospective study, the generalizability of its results is limited. The potential use of prospective design and multicenter data in future studies is expected to improve model performance and further validate the robustness and generalizability of the model.

Conclusions

In summary, we successfully developed an interpretable ML model to predict the risk of developing traumatic cerebral infarction based on clinical data easily extracted from a case system. The final LR model has excellent predictive power in both internal and external validation. In addition, with the SHAP values, we provide a personalized risk assessment, which provides important decision support to physicians. This effective computer-assisted approach can help frontline clinicians identify risks and intervene early, promoting individualized management and thus reducing PTCI events.

Data Sharing Statement

Research data are not available at this time.

Ethical Approval and Consent to Participate

The study protocol was approved by the Ethics Committee of Fuzhou University Affiliated Provincial Hospital (K2024-06-057). This study was conducted in accordance with the Declaration of Helsinki. The need for written informed consent was waived by the Ethics Committee of Fuzhou University Affiliated Provincial Hospital due to the non-interventional design of the study. Additionally, this research exclusively utilized previously collected medical record information from which all personally identifiable information had been removed, ensuring no risk to the subjects and no adverse effects on their rights and health.

Acknowledgments

The authors would like to express their gratitude to the following institutions: Fujian Provincial Key Laboratory of Medical Big Data Engineering and Fujian Provincial Key Laboratory of Emergency Medicine.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by Major scientific research project for middle-aged and young people in health in Fujian Province (2021ZQNZD008).

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Pavlovic D, Pekic S, Stojanovic M, Popovic V. Traumatic brain injury: neuropathological, neurocognitive and neurobehavioral sequelae. *Pituitary*. 2019;22:270–282. doi:10.1007/s11102-019-00957-9
2. Latronico N, Piva S, Fagoni N, et al. Impact of a posttraumatic cerebral infarction on outcome in patients with TBI: the Italian multicenter cohort INCEPT study. *Crit Care Lond Engl*. 2020;24:33. doi:10.1186/s13054-020-2746-5
3. Server A, Dullerud R, Haakonsen M, et al. Post-traumatic cerebral infarction. Neuroimaging findings, etiology and outcome. *Acta Radiol Stockh Swed*. 2001;42:254–260. doi:10.1080/028418501127346792
4. Zhi-Ling C, Qi L, Jun-Yong Y, Bang-Qing Y. The prevalence and risk factors of posttraumatic cerebral infarction in patients with traumatic brain injury: a systematic review and meta-analysis. *Bioengineered*. 2022;13:11706–11717. doi:10.1080/21655979.2022.2070999
5. Liu S, Wan X, Wang S, et al. Posttraumatic cerebral infarction in severe traumatic brain injury: characteristics, risk factors and potential mechanisms. *Acta Neurochir*. 2015;157:1697–1704. doi:10.1007/s00701-015-2559-5
6. Benke T, Dazinger F, Pechlaner R, et al. Lesion topography of posterior cerebral artery infarcts. *J Neurol Sci*. 2021;428:117585. doi:10.1016/j.jns.2021.117585
7. Wijesundera C, Vingrys AJ, Wijeratne T, Crewther SG. Acquired visual deficits independent of lesion site in acute stroke. *Front Neurol*. 2020;11:705. doi:10.3389/fneur.2020.00705
8. Sav A, Rotondo F, Syro LV, et al. Pituitary pathology in traumatic brain injury: a review. *Pituitary*. 2019;22(3):201–211. doi:10.1007/s11102-019-00958-8
9. Heo J, Yoon JG, Park H, et al. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke*. 2019;50:1263–1265. doi:10.1161/STROKEAHA.118.024293
10. Ota R, Yamashita F. Application of machine learning techniques to the analysis and prediction of drug pharmacokinetics. *J Control Release off J Control Release Soc*. 2022;352:961–969. doi:10.1016/j.jconrel.2022.11.014
11. Deo RC, Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593
12. Mirvis SE, Wolf AL, Numaguchi Y, et al. Posttraumatic cerebral infarction diagnosed by CT: prevalence, origin, and outcome. *AJR Am J Roentgenol*. 1990;154:1293–1298. doi:10.2214/ajr.154.6.2110744
13. Wardlaw JM, Mielke O. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment--systematic review. *Radiology*. 2005;235:444–453. doi:10.1148/radiol.2352040262
14. Marino R, Gasparotti R, Pinelli L, et al. Posttraumatic cerebral infarction in patients with moderate or severe head trauma. *Neurology*. 2006;67:1165–1171. doi:10.1212/01.wnl.0000238081.35281.b5
15. Tian H-L, Geng Z, Cui Y-H, et al. Risk factors for posttraumatic cerebral infarction in patients with moderate or severe head trauma. *Neurosurg Rev*. 2008;31:431–436. doi:10.1007/s10143-008-0153-5
16. Hu X, Tian J, Xie J, et al. Predictive role of shock index in the early formation of cerebral infarction in patients with TBI and cerebral herniation. *Front Neurol*. 2022;13:956039. doi:10.3389/fneur.2022.956039
17. Wu Y-G, Chao Y, Gao G, et al. Risk factors for cerebral infarction after moderate or severe traumatic brain injury. *Ther Clin Risk Manag*. 2021;17:433–440. doi:10.2147/TCRM.S309662
18. Tyrrell DJ, Goldstein DR. Ageing and atherosclerosis: vascular intrinsic and extrinsic factors and potential role of IL-6. *Nat Rev Cardiol*. 2021;18:58–68. doi:10.1038/s41569-020-0431-7
19. Schönberger M, Ponsford J, Reutens D, et al. The relationship between age, injury severity, and MRI findings after traumatic brain injury. *J Neurotrauma*. 2009;26(12):2157–2167. doi:10.1089/neu.2009.0939
20. Czosnyka M, Balestreri M, Steiner L, et al. Age, intracranial pressure, autoregulation, and outcome after brain trauma. *J Neurosurg*. 2005;102:450–454. doi:10.3171/jns.2005.102.3.0450

21. Sun -Y-Y, Zhu H-J, Zhao R-Y, et al. Remote ischemic conditioning attenuates oxidative stress and inflammation via the Nrf2/HO-1 pathway in MCAO mice. *Redox Biol*. 2023;66:102852. doi:10.1016/j.redox.2023.102852
22. Wang M, Wu J, Jiao H, et al. Enterocyte synthesizes and secrets uric acid as antioxidant to protect against oxidative stress via the involvement of Nrf pathway. *Free Radic Biol Med*. 2022;179:95–108. doi:10.1016/j.freeradbiomed.2021.12.307
23. Amaro S, Laredo C, Renú A, et al. Uric acid therapy prevents early ischemic stroke progression: a tertiary analysis of the URICO-ICTUS trial (Efficacy study of combined treatment with uric acid and r-tPA in acute ischemic stroke). *Stroke*. 2016;47:2874–2876. doi:10.1161/STROKEAHA.116.014672
24. Hoffman AN, Taylor AN. Stress reactivity after traumatic brain injury: implications for comorbid post-traumatic stress disorder. *Behav Pharmacol*. 2019;30:115–121. doi:10.1097/FBP.0000000000000461
25. Poznyak A, Grechko AV, Poggio P, et al. The diabetes mellitus-atherosclerosis connection: the role of lipid and glucose metabolism and chronic inflammation. *Int J Mol Sci*. 2020;21:1835. doi:10.3390/ijms21051835
26. Shi J, Dong B, Mao Y, et al. Review: traumatic brain injury and hyperglycemia, a potentially modifiable risk factor. *Oncotarget*. 2016;7:71052–71061. doi:10.18632/oncotarget.11958
27. Su T-M, Lan C-M, Lee T-H, et al. Posttraumatic cerebral infarction after decompressive craniectomy for traumatic brain injury: incidence, risk factors and outcome. *Turk Neurosurg*. 2017. doi:10.5137/1019-5149.JTN.20761-17.1
28. Sang Y, Roest M, de Laat B, et al. Interplay between platelets and coagulation. *Blood Rev*. 2021;46:100733. doi:10.1016/j.blre.2020.100733
29. Ranucci M, Ranucci M, Laddomada T, et al. Plasma viscosity, functional fibrinogen, and platelet reactivity in vascular surgery patients. *Clin Hemorheol Microcirc*. 2015;61:417–427. doi:10.3233/CH-141866
30. Slavin KV, Vannemreddy P. Neuromodulation for post-SAH cerebral vasospasm. *Neuromodulation J Int Neuromodulation Soc*. 2022;25:1444–1445. doi:10.1016/j.neurom.2022.08.460
31. Isola C, Evain J-N, Francony G, et al. Cerebral vasospasm in children with subarachnoid hemorrhage: frequency, diagnosis, and therapeutic management. *Neurocrit Care*. 2022;36:868–875. doi:10.1007/s12028-021-01388-w
32. Langham J, Goldfrad C, Teasdale G, et al. Calcium channel blockers for acute traumatic brain injury. *Cochrane Database Syst Rev*. 2003. doi:10.1002/14651858.CD000565
33. Xu G-Z, Wang M-D, Liu K-G, et al. A meta-analysis of treating acute traumatic brain injury with calcium channel blockers. *Brain Res Bull*. 2013;99:41–47. doi:10.1016/j.brainresbull.2013.09.011