

Development and Validation of a Machine Learning Model Integrating Stress Hyperglycemia for Predicting Early Neurological Deterioration in Acute Mild-to-Moderate Ischemic Stroke

Zhihui Chen, Zhixin Huang*, Kaili Huang

Department of Neurology, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, Guangzhou 510317, Guangdong, China

*Corresponding author: Zhixin Huang, huangjamie@126.com

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: *Objective:* To investigate the predictive value of stress hyperglycemia and various machine learning models for early neurological deterioration (END) in patients with acute mild-to-moderate ischemic stroke, and to provide a quantitative reference for early clinical identification of high-risk patients and formulation of individualized intervention strategies. *Methods:* In this retrospective study, a total of 1,169 patients with acute mild-to-moderate ischemic stroke who met the inclusion criteria were enrolled from Guangdong Second Provincial General Hospital between January 2017 and June 2023. After applying the predefined exclusion criteria, 933 eligible subjects were finally included. *Results:* Among the 933 enrolled patients, the END incidence was 18.97%. Baseline data showed that the END group had higher body mass index (BMI), admission blood pressure, baseline National Institutes of Health Stroke Scale (NIHSS) score, admission blood glucose, glycated hemoglobin, and stress hyperglycemia ratio (SHR), while Alberta Stroke Program Early CT Score (ASPECTS) and Activities of Daily Living (ADL) scores were lower. Stress hyperglycemia was also more frequent in the END group. A history of hypertension was significantly more common in the END group, whereas the difference in diabetes mellitus history was not statistically significant ($P = 0.064$). Four machine learning prediction models were constructed based on the dataset divided at a 7:3 ratio. XGBoost may have potential utility for sensitivity-oriented screening, although the current findings do not demonstrate its overall superiority over the other models. *Conclusion:* Four machine learning models integrating stress hyperglycemia can effectively predict the risk of early neurological deterioration in patients with acute mild-to-moderate ischemic stroke. Among the evaluated models, XGBoost showed potential utility for sensitivity-oriented screening and may support risk stratification.

Keywords: Stress hyperglycemia; Acute mild-to-moderate ischemic stroke; Early neurological deterioration; Machine learning; Prediction model

Online publication: October 19, 2026

1. Introduction

Acute ischemic stroke is a leading cause of mortality and disability among Chinese residents, with mild-to-moderate ischemic stroke accounting for nearly half of all ischemic stroke cases ^[1]. Although the overall prognosis for mild-to-moderate ischemic stroke is generally considered favorable, 10% to 30% of patients nevertheless experience early neurological deterioration within days of onset, resulting in poor outcomes, disability, or death ^[2-5]. Early identification of high-risk individuals and timely clinical intervention are therefore critical to improving prognosis. Recent studies have shown that stress hyperglycemia is associated with poor prognosis in acute ischemic stroke ^[6-8]. Even in non-diabetic patients, elevated post-onset blood glucose levels increase the risk of neurological deterioration. However, few studies have integrated stress hyperglycemia into prediction models for END risk assessment in this population. Machine learning algorithms can efficiently integrate multidimensional clinical features, improving the accuracy and efficiency of disease risk prediction, and have been increasingly applied to cerebrovascular prognostic research. This study aimed to develop and validate machine learning prediction models integrating stress hyperglycemia indicators based on clinical data from patients with acute mild-to-moderate ischemic stroke, and to evaluate the predictive performance of different models to support early clinical identification of high-risk patients.

2. Materials and methods

2.1. Study design and population

A retrospective study design was adopted. A total of 1,169 patients with acute mild-to-moderate ischemic stroke who met the inclusion criteria were enrolled from a hospital between January 2017 and June 2023. The study protocol was approved by the institutional ethics committee (No. 2025-KY-KZ-310-01) and was conducted in accordance with the Declaration of Helsinki. According to the predefined exclusion criteria, 4 patients with a pre-stroke modified Rankin Scale (mRS) score >2, 28 with recent severe infection, 25 with thyroid disorders, 8 with hematological diseases (e.g., moderate-to-severe anemia), 12 with severe cardiac, hepatic, or renal dysfunction, 5 with malignant tumors, and 154 with incomplete clinical data were sequentially excluded. A total of 933 eligible subjects were finally included. Patients were divided into END and non-END groups according to the occurrence of END after onset, with 177 cases (18.97%) and 756 cases, respectively. The baseline characteristics of the two groups are summarized in **Table 1**.

The 933 included patients ranged in age from 18 to 99 years, with a mean of 59.28 ± 11.85 years and a median (interquartile range) of 58 (51–67) years. In the non-END group ($n = 756$), ages ranged from 18 to 99 years, with a mean of 59.60 ± 11.96 years and a median of 59 (51–68) years. In the END group ($n = 177$), ages ranged from 22 to 85 years, with a mean of 57.93 ± 11.34 years and a median of 56 (50–66) years. The difference in age between the two groups was not statistically significant ($P > 0.05$).

2.1.1. Inclusion criteria

- (1) Diagnosis of acute mild-to-moderate ischemic stroke conforming to the relevant criteria in the Chinese Guidelines for the Diagnosis and Treatment of Acute Ischemic Stroke ^[9], confirmed by cranial CT or MRI;
- (2) Admission within 72 hours of onset;
- (3) National Institutes of Health Stroke Scale (NIHSS) score ≤ 15 after onset;
- (4) Informed consent signed by the patient or family members.

2.1.2. Exclusion criteria

- (1) Pre-stroke modified Rankin Scale (mRS) score ≥ 3 , indicating pre-existing neurological deficits that could affect prognostic assessment;
- (2) Severe infection within 4 weeks before onset or ongoing immunosuppressive therapy;
- (3) Thyroid dysfunction, active bleeding, moderate-to-severe anemia, or other hematological diseases;
- (4) Severe cardiac, hepatic, or renal insufficiency;
- (5) Malignant tumors or other diseases with an expected survival of less than 1 year;
- (6) Incomplete clinical or laboratory data precluding the required analyses.

2.2. Methods

2.2.1. Data collection

2.2.1.1. Baseline characteristic collection

A total of 62 research indicators were collected and categorized into 11 major dimensions, comprehensively covering baseline data, clinical assessments, laboratory tests, and personal history. The specific classifications were as follows:

- (1) Demographic characteristics: age, sex, onset time, height, weight, and body mass index (BMI) (6 items).
- (2) Vital signs and functional scores: admission systolic BP, admission diastolic blood pressure, and Activities of Daily Living (ADL) score (3 items).
- (3) Past medical history: history of hypertension, diabetes mellitus, atrial fibrillation, prior cerebral infarction, and coronary heart disease (5 categorical items).
- (4) Medication history: glucose-lowering medication use, antihypertensive medication use, anticoagulant use, antiplatelet use, and statin use (5 categorical items).
- (5) Lifestyle habits: smoking history and drinking history (2 categorical items).
- (6) Neurological and treatment-related characteristics: admission National Institutes of Health Stroke Scale (NIHSS) score, Alberta Stroke Program Early CT Score (ASPECTS), infarct location, Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification, and treatment modality (5 items).
- (7) Routine blood tests: white blood cell count, absolute neutrophil count, absolute monocyte count, absolute eosinophil count, absolute basophil count, absolute lymphocyte count, red blood cell count, hemoglobin, and platelet count (9 items).
- (8) Thyroid function: thyroid-stimulating hormone, free T3 and free T4 (3 items).
- (9) Coagulation function: prothrombin time, thrombin time, international normalized ratio, D-dimer, activated partial thromboplastin time, and fibrin degradation products (6 items).
- (10) Blood biochemistry: total cholesterol, triglycerides, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, alanine aminotransferase, aspartate aminotransferase, homocysteine, C-reactive protein, creatinine, uric acid, admission blood glucose, creatine kinase, creatine kinase-MB, total protein, and albumin (15 items).
- (11) Glucose metabolism indicators: glycated hemoglobin, stress hyperglycemia ratio, and stress hyperglycemia (3 items).

These indicators collectively cover patients' baseline personal information, past medical history and medication use, lifestyle habits, neurological functional status, core organ function, and various biochemical, coagulation, and metabolic laboratory tests. The comprehensive data dimensions and complete content provide sufficient support for subsequent statistical analyses.

2.2.1.2. Definition of key indicators

Stress hyperglycemia (SHG) was defined as follows: for patients without a history of diabetes mellitus, admission random blood glucose ≥ 7.8 mmol/L; for patients with a history of diabetes mellitus, blood glucose elevated by ≥ 2.5 mmol/L above the baseline level^[10-14]. SHG was coded as 1 if present and 0 if absent.

The stress hyperglycemia ratio (SHR) was calculated using the following formula:

$$\text{SHR} = \frac{\text{Admission blood glucose}}{1.59 \times \text{HbA1c} - 2.59}$$

where SHR is calculated as admission random blood glucose divided by the estimated average glucose derived from glycated hemoglobin, i.e., $1.59 \times \text{HbA1c} - 2.59$. SHR was analyzed as a continuous indicator and was not used to define stress hyperglycemia. Stress hyperglycemia classification was based only on the admission glucose criteria described above. All laboratory tests were performed by professional technicians using standardized equipment. All clinical assessments were independently scored by two neurologists who had received unified training. In cases of disagreement, a third senior neurologist performed a review to confirm the final score, ensuring the accuracy and reliability of all data included in the study. All raw data were compiled by researchers, entered into a dedicated database, verified through double data entry, and locked for subsequent model development and validation analysis.

2.2.1.3. Outcome definition

Early neurological deterioration was defined as an increase in the total NIHSS score of ≥ 2 points from baseline, or an increase of ≥ 1 point in the motor item score, at any repeated NIHSS assessment within 72 hours of admission.

2.3. Data preprocessing and feature engineering

The collected raw data were cleaned and organized. Features with missing proportions exceeding 30% were excluded. The remaining missing data were imputed using multiple imputation methods. All continuous features were standardized to eliminate the influence of different units and scales on model training. Categorical features were transformed using one-hot encoding. All subjects were randomly divided into a training set and a test set at a 7:3 ratio. The training set was used for model training and feature selection, while the test set was used for independent performance evaluation. The least absolute shrinkage and selection operator (LASSO) algorithm was applied to the 62 original features in the training set to eliminate redundant features with low contribution, retain core features with high predictive value, reduce overfitting risk, and improve model efficiency.

2.4. Model construction and validation

Four machine learning prediction models were constructed: Logistic regression, support vector machine (SVM), random forest, and extreme gradient boosting (XGBoost). Model training and parameter tuning were completed on the training set, with Logistic regression serving as a baseline reference. The area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were used to evaluate model performance in both the training and test sets. Calibration curves combined with the Hosmer-Lemeshow test were used to assess calibration. Feature

importance analysis was performed to determine the contribution of each feature. Decision curve analysis (DCA) was used to calculate the net benefit of the models at various threshold probabilities.

2.5. Statistical analysis

All statistical analyses were performed using R language and Python software. Continuous variables following a normal distribution were expressed as mean $\pm$ standard deviation, and between-group comparisons using independent sample *t*-tests. Non-normally distributed variables were expressed as median (interquartile range), and between-group comparisons using rank-sum tests. Categorical variables were expressed as counts (percentages), and between-group comparisons using chi-square or Fisher's exact tests. All tests were two-sided, and $P < 0.05$ was considered statistically significant.

3. Results

3.1. Baseline characteristics of the study population

A total of 933 subjects were included, of whom 177 developed END (18.97%) and 756 did not. Compared with the non-END group, the END group had significantly higher height, weight, BMI, admission systolic and diastolic blood pressures, admission NIHSS score, white blood cell count, absolute neutrophil count, red blood cell count, hemoglobin, total protein, admission blood glucose, glycated hemoglobin, and SHR, while admission ADL score, ASPECTS, and absolute eosinophil count were significantly lower (all $P < 0.05$). A history of hypertension and antihypertensive medication use were more common in the END group, and stress hyperglycemia was also more frequent (all $P < 0.05$). The distributions of TOAST classification and treatment modality differed significantly between groups (both $P < 0.001$). By contrast, the difference in diabetes mellitus history did not reach statistical significance ($P = 0.064$), and infarct location also showed no significant between-group comparisons ($P = 0.075$). The remaining indicators showed no statistically significant differences ($P > 0.05$).

Table 1. Baseline characteristics of the study population

Variable	Total (n=933)	Non-END (n=756)	END (n=177)	Test Statistic	P value
Demographic characteristics					
Sex (male)	685 (73.4%)	556 (73.5%)	129 (72.9%)	0.03	0.857
Age(year)	58.00 (51.00, 67.00)	59.00 (51.00, 68.00)	56.00 (50.00, 66.00)	1.55	0.121
Onset time (h)	13.00 (4.00, 31.00)	13.00 (4.00, 30.65)	12.00 (4.50, 32.00)	-0.41	0.685
Height (m)	1.65 (1.60, 1.70)	1.65 (1.60, 1.70)	1.66 (1.60, 1.70)	-1.98	0.047
Weight (kg)	65.00 (60.00, 75.00)	65.00 (60.00, 74.00)	68.00 (62.00, 75.00)	-3.12	0.002
BMI (kg/m ²)	24.22 (22.38, 26.57)	24.22 (22.09, 26.42)	24.49 (23.44, 27.34)	-2.99	0.003
Admission systolic BP(mmHg)	151.00 (135.00, 167.00)	149.50 (134.00, 165.00)	157.00 (139.00, 177.00)	-3.79	<0.001
Admission diastolic BP(mmHg)	89.00 (80.00, 100.00)	88.00 (79.00, 100.00)	92.00 (85.00, 102.00)	-3.27	0.001
Admission ADL score	75.00 (50.00, 95.00)	75.00 (55.00, 95.00)	55.00 (35.00, 85.00)	7.23	<0.001
History of hypertension	587 (62.9%)	464 (61.4%)	123 (69.5%)	4.05	0.044
History of diabetes mellitus	229 (24.5%)	176 (23.3%)	53 (29.9%)	3.44	0.064
History of atrial fibrillation	11 (1.2%)	9 (1.2%)	2 (1.1%)	—	>0.999

Variable	Total (n=933)	Non-END (n=756)	END (n=177)	Test Statistic	P value
Demographic characteristics					
Prior cerebral infarction	129 (13.8%)	107 (14.2%)	22 (12.4%)	0.36	0.550
History of coronary heart disease	54 (5.8%)	45 (6.0%)	9 (5.1%)	0.20	0.656
History of antihypertensive medication use	422 (45.2%)	327 (43.3%)	95 (53.7%)	6.28	0.012
History of glucose-lowering medication use	177 (19.0%)	137 (18.1%)	40 (22.6%)	1.87	0.171
History of anticoagulant use	7 (0.8%)	7 (0.9%)	0 (0.0%)	—	0.358
History of antiplatelet use	90 (9.6%)	74 (9.8%)	16 (9.0%)	0.09	0.761
History of statin use	87 (9.3%)	72 (9.5%)	15 (8.5%)	0.19	0.666
Smoking history	461 (49.4%)	374 (49.5%)	87 (49.2%)	0.01	0.939
Drinking history	272 (29.2%)	221 (29.2%)	51 (28.8%)	0.01	0.912
Admission NIHSS score	3.00 (2.00, 5.00)	3.00 (2.00, 4.00)	4.00 (2.00, 6.00)	-4.56	<0.001
ASPECTS	9.00 (8.00, 9.00)	9.00 (8.00, 9.00)	8.00 (7.00, 9.00)	8.07	<0.001
Infarct location				5.18	0.075
1	641 (68.7%)	531 (70.2%)	110 (62.1%)		
2	187 (20.0%)	147 (19.4%)	40 (22.6%)		
3	105 (11.3%)	78 (10.3%)	27 (15.3%)		
TOAST classification				—	<0.001
1	434 (46.5%)	321 (42.5%)	113 (63.8%)		
2	26 (2.8%)	21 (2.8%)	5 (2.8%)		
3	451 (48.3%)	394 (52.1%)	57 (32.2%)		
4	17 (1.8%)	15 (2.0%)	2 (1.1%)		
5	5 (0.5%)	5 (0.7%)	0 (0.0%)		
Treatment modality				59.57	<0.001
1	637 (68.3%)	539 (71.3%)	98 (55.4%)		
2	127 (13.6%)	75 (9.9%)	52 (29.4%)		
3	27 (2.9%)	19 (2.5%)	8 (4.5%)		
4	41 (4.4%)	29 (3.8%)	12 (6.8%)		
5	101 (10.8%)	94 (12.4%)	7 (4.0%)		
White blood cell count	8.04 (6.60, 9.65)	7.96 (6.60, 9.44)	8.54 (6.88, 10.09)	-2.09	0.036
Absolute neutrophil count	5.30 (4.10, 6.97)	5.26 (4.09, 6.75)	5.91 (4.45, 7.64)	-3.00	0.003
Absolute monocyte count	0.46 (0.36, 0.58)	0.46 (0.36, 0.58)	0.46 (0.36, 0.58)	0.04	0.965
Absolute eosinophil count	0.10 (0.04, 0.18)	0.10 (0.05, 0.20)	0.08 (0.04, 0.14)	2.66	0.008
Absolute basophil count	0.03 (0.02, 0.04)	0.03 (0.02, 0.04)	0.03 (0.02, 0.04)	-0.93	0.351
Absolute lymphocyte count	1.81 (1.34, 2.31)	1.82 (1.34, 2.33)	1.79 (1.34, 2.22)	1.26	0.209
Red blood cell count	4.79 (4.41, 5.18)	4.78 (4.40, 5.13)	4.94 (4.55, 5.29)	-2.89	0.004
Hemoglobin	144.00 (131.00, 155.00)	143.00 (131.00, 154.00)	148.00 (135.00, 157.00)	-2.19	0.029
Platelet count	238.00 (201.00, 281.00)	237.50 (200.00, 278.50)	243.00 (203.00, 288.00)	-1.33	0.185
Thyroid-stimulating hormone	1.39 (0.87, 2.06)	1.44 (0.87, 2.09)	1.26 (0.87, 1.74)	1.94	0.052
Free triiodothyronine	4.39 (3.97, 4.86)	4.39 (3.97, 4.83)	4.43 (3.97, 5.07)	-0.93	0.353
Free thyroxine	15.98 (14.60, 17.30)	15.95 (14.60, 17.21)	16.04 (14.64, 17.63)	-0.61	0.543
Prothrombin time	12.90 (12.40, 13.50)	12.90 (12.40, 13.50)	12.90 (12.60, 13.40)	-0.74	0.458

Variable	Total (n=933)	Non-END (n=756)	END (n=177)	Test Statistic	P value
Demographic characteristics					
Thrombin time	16.20 (15.60, 16.80)	16.20 (15.60, 16.80)	16.20 (15.60, 16.90)	-0.03	0.978
International normalized ratio	0.99 (0.94, 1.03)	0.99 (0.94, 1.04)	0.99 (0.95, 1.03)	-0.57	0.566
D-dimer	0.29 (0.22, 0.49)	0.29 (0.22, 0.49)	0.30 (0.22, 0.53)	-1.12	0.261
Activated partial thromboplastin time	35.70 (32.80, 38.50)	35.70 (32.80, 38.65)	35.40 (32.80, 37.50)	1.00	0.319
Fibrin degradation products (mg/L)	13.80 (12.00, 15.80)	13.80 (11.90, 15.85)	13.80 (12.10, 15.80)	-0.39	0.699
Total cholesterol	4.95 (4.28, 5.66)	4.95 (4.26, 5.62)	5.00 (4.40, 5.78)	-1.56	0.119
Triglycerides	1.44 (1.05, 2.17)	1.43 (1.03, 2.17)	1.48 (1.12, 2.22)	-1.70	0.089
High-density lipoprotein cholesterol	1.09 (0.92, 1.30)	1.09 (0.92, 1.30)	1.09 (0.92, 1.29)	0.18	0.859
Low-density lipoprotein cholesterol	2.98 (2.40, 3.62)	2.98 (2.40, 3.62)	2.98 (2.43, 3.72)	-0.43	0.667
Alanine aminotransferase	21.00 (16.00, 30.00)	21.00 (16.00, 29.00)	23.00 (17.00, 30.00)	-1.55	0.122
Aspartate aminotransferase	19.00 (16.00, 23.00)	19.00 (16.00, 23.00)	19.00 (16.00, 23.00)	0.46	0.645
Homocysteine	12.70 (10.30, 15.30)	12.80 (10.23, 15.30)	12.57 (10.30, 15.32)	0.25	0.799
C-reactive protein	2.80 (0.30, 6.74)	2.72 (0.20, 6.70)	3.30 (0.50, 7.00)	-1.10	0.269
Creatinine	83.00 (68.00, 101.00)	83.00 (68.10, 101.00)	81.00 (67.10, 99.00)	0.65	0.514
Uric acid	370.00 (304.00, 443.00)	371.00 (304.00, 440.50)	368.00 (301.00, 456.00)	-0.39	0.693
Admission blood glucose	6.38 (5.27, 8.59)	6.23 (5.21, 8.24)	7.06 (5.73, 10.02)	-4.06	<0.001
Creatine kinase	94.00 (68.00, 134.00)	94.00 (68.00, 132.00)	94.00 (74.00, 145.00)	-0.97	0.330
Creatine kinase-MB	8.20 (3.41, 12.30)	8.20 (3.50, 12.55)	7.20 (3.10, 10.90)	1.79	0.074
Total protein	71.80 (66.90, 76.90)	71.80 (66.60, 76.70)	72.90 (68.10, 78.60)	-2.70	0.007
Albumin	40.60 (37.60, 43.80)	40.60 (37.45, 43.60)	41.00 (38.50, 44.10)	-1.60	0.110
Glycated hemoglobin	6.00 (5.60, 6.80)	6.00 (5.60, 6.70)	6.10 (5.70, 7.30)	-2.03	0.042
Stress hyperglycemia	192 (20.6%)	145 (19.2%)	47 (26.6%)	4.77	0.029
Stress hyperglycemia ratio	0.91 (0.77, 1.09)	0.89 (0.76, 1.08)	0.96 (0.83, 1.13)	-3.05	0.002
¹ Median (Q1, Q3); n (%)					
² Wilcoxon rank-sum test (standardized Z); Pearson's Chi-squared test; Fisher's exact test					

3.2. Model performance comparison

All 933 subjects were randomly divided into a training set and a test set at a 7:3 ratio. The test set included 280 patients, of whom 53 were END-positive. Four machine learning models were constructed based on admission baseline indicators. ROC curves for each model in the test set are shown in **Figure 1**.

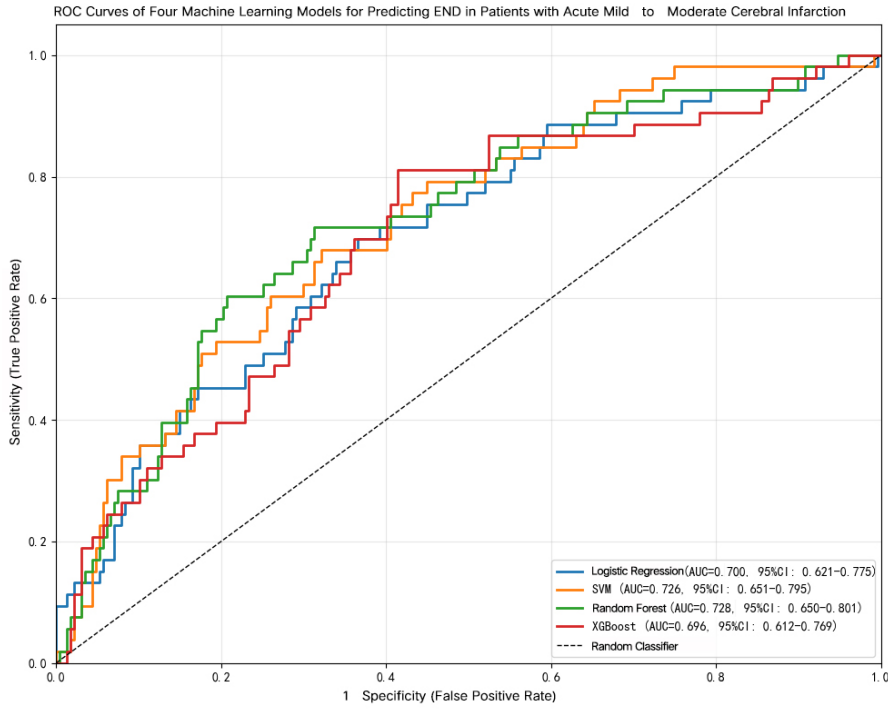


Figure 1. ROC curves of the four models on the test set

The AUC values, in descending order, were: random forest, 0.728 (95% CI: 0.650–0.801); SVM, 0.726 (95% CI: 0.651–0.795); Logistic regression, 0.700 (95% CI: 0.621–0.775); and XGBoost, 0.696 (95% CI: 0.612–0.769). All four models demonstrated moderate predictive performance for END.

Using the XGBoost model as the reference for the DeLong test, pairwise comparisons of AUC between the XGBoost model and the Logistic regression, SVM, and random forest models showed no statistically significant differences ($P = 0.9105, 0.3079, \text{ and } 0.1422$, respectively; all $P > 0.05$).

After determining the optimal cutoff values based on the Youden index, sensitivity, specificity, accuracy, positive predictive value, negative predictive value, and F1 score were calculated. The random forest model achieved the highest F1 score (0.469) and the highest specificity (0.687), with a sensitivity of 0.717 and an accuracy of 0.693. The XGBoost model demonstrated the highest sensitivity (0.811), making it potentially useful for initial screening of high-risk populations, although its specificity was relatively low (0.586). The Logistic regression model had a specificity of 0.634, while the SVM model exhibited relatively balanced performance across the evaluated indicators. The negative predictive values of all four models exceeded 0.90, indicating potential clinical utility for excluding low-risk patients.

3.3. Feature importance analysis based on SHapley Additive exPlanations (SHAP)

To explore the contribution of each clinical indicator to the prediction of early neurological deterioration (END) in patients with acute mild-to-moderate ischemic stroke, this study applied the SHapley Additive exPlanations (SHAP) algorithm to the XGBoost prediction model for feature analysis, quantifying the contribution of each variable to the risk of END occurrence. The features, ranked in descending order by mean absolute SHAP value, were: ASPECTS (0.5806), BMI (0.3898), admission diastolic blood pressure (0.3509), admission systolic BP (0.3287), admission NIHSS score (0.3247), age (0.2900), stress hyperglycemia ratio (SHR, 0.2608), onset time (0.2081), admission blood glucose (0.1899), and sex (0.0314), as shown in **Figure 2**.

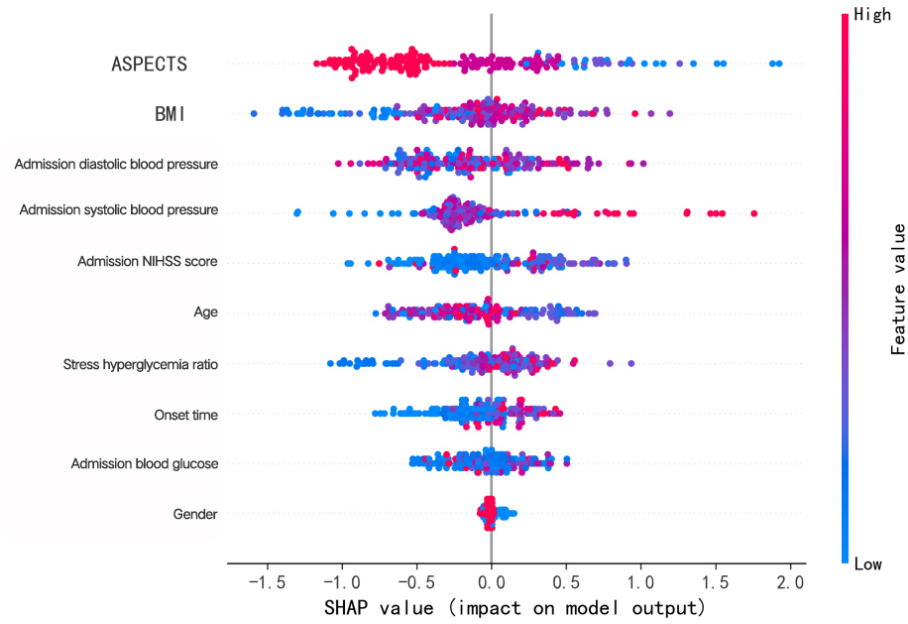


Figure 2. SHAP summary plot of the XGBoost model

As shown in **Figure 2**, the magnitude of predictive contribution of each feature and the direction of influence of feature values on the prediction results can be observed. Among all features, ASPECTS was the strongest predictor of END risk, with lower scores indicating larger infarct extent and higher risk of END. BMI, admission blood pressure, and baseline neurological deficit severity (admission NIHSS score) were also key predictors of END. SHAP values demonstrated that SHR ranked seventh with respect to predictive contribution. The predominance of positive SHAP values indicated that higher SHR was linked to increased predicted probability of END within the model. It should be noted that SHAP quantifies feature-level contribution to model-based prediction, which does not constitute evidence for SHR as an independent risk factor derived from conventional regression.

The nonlinear positive correlation between SHR and the predicted probability is shown in **Figure 3**.

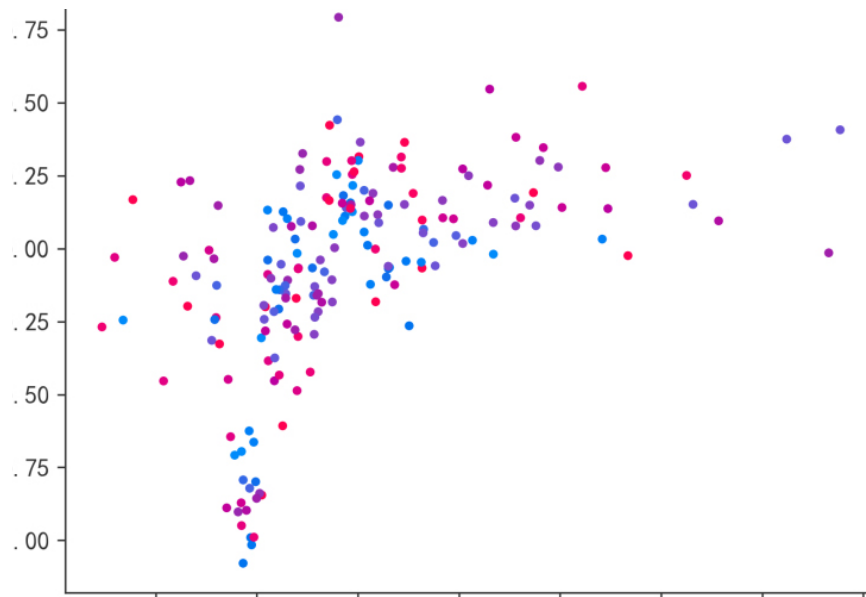


Figure 3. SHAP dependence plot for SHR

As shown in **Figure 3**, the SHAP dependence plot for the stress hyperglycemia ratio intuitively reflects the nonlinear relationship between SHR and the predicted risk of END: as SHR values increased, the corresponding SHAP values showed an overall upward trend, and the predicted probability of END gradually increased, suggesting that stress hyperglycemia may contribute to END risk prediction. Sex had the lowest predictive contribution, with a mean absolute SHAP value of only 0.0314, indicating a relatively weak impact on model predictive performance.

3.4. Decision curve analysis

In this study, Logistic regression and XGBoost prediction models were selected for decision curve analysis, with the risk threshold range limited to the clinically common interval of 0 to 0.6 in order to evaluate the clinical net benefit of the two models in screening high-risk populations for early neurological deterioration (END) in patients with acute mild-to-moderate ischemic stroke.

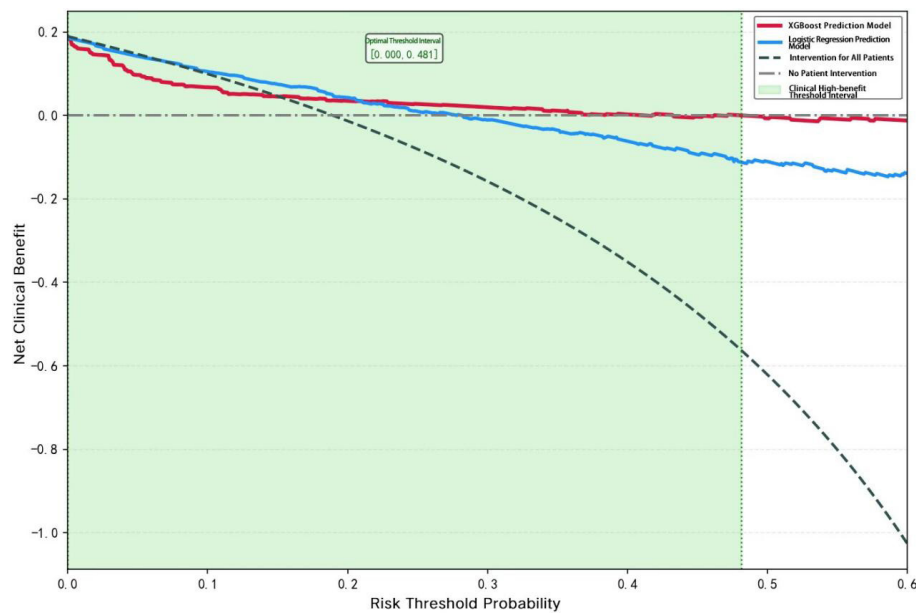


Figure 4. Decision curve analysis

As shown in **Figure 4**, the net benefit curves did not remain above the “treat all” and “treat none” reference strategies across the entire 0 to 0.6 threshold range. The Logistic regression curve approached and crossed the zero-net-benefit line at approximately 0.27, whereas the XGBoost curve retained positive net benefit over a wider interval, extending to approximately 0.481, which corresponds to the threshold interval marked in **Figure 4**. Beyond these ranges, net benefit declined, and neither model maintained a positive advantage throughout the full evaluated interval. Therefore, the DCA findings support potential clinical utility within selected threshold ranges rather than across all thresholds from 0 to 0.6. DeLong comparisons among the four models showed no statistically significant differences in AUC (all $P > 0.05$). The negative predictive values of both Logistic regression and XGBoost exceeded 0.90, suggesting potential value for excluding low-risk END patients.

4. Discussion

4.1. Analysis of baseline risk factors for END in patients with acute mild-to-moderate ischemic stroke

A total of 933 patients were included, with an END incidence of 18.97%, consistent with the 17%–20% range reported in recent studies. Between-group comparisons using showed that the END group had significantly higher BMI, admission blood pressure, admission NIHSS score, admission blood glucose, glycated hemoglobin, white blood cell count, neutrophil count, red blood cell count, hemoglobin, total protein, and SHR, and significantly lower ASPECTS, ADL score, and eosinophil count. Height and weight were also higher in the END group, while stress hyperglycemia was more frequent. A history of hypertension and antihypertensive medication use were significantly more common in the END group. In contrast, diabetes mellitus history was numerically more frequent but did not reach statistical significance ($P = 0.064$), and infarct location did not differ significantly between groups ($P = 0.075$). TOAST classification and treatment modality showed significant differences in distribution rather than ordinal “higher” or “lower” scores.

Obesity and hypertension are classic modifiable risk factors for stroke. Elevated BMI may aggravate vascular endothelial dysfunction and low-grade systemic inflammation, while abnormally elevated admission blood pressure may contribute to hemodynamic instability around the infarct and worsen ischemic penumbral injury. Higher baseline NIHSS and lower ASPECTS indicate greater initial neurological deficit and ischemic burden. Elevated admission blood glucose, glycated hemoglobin, stress hyperglycemia, and SHR in the END group suggest that acute metabolic stress is closely associated with early neurological worsening. Differences in leukocyte and neutrophil counts are also consistent with an inflammatory contribution. The marked between-group comparisons in treatment modality ($P < 0.001$) requires particular caution because treatment selection is related to baseline severity, onset-to-admission time, imaging findings, and treatment eligibility and may itself influence the early neurological course. Treatment modality was included among the candidate variables during feature selection, but residual confounding by indication cannot be fully excluded in this retrospective analysis. Future external validation should further stratify or adjust for treatment categories when assessing END risk.

4.2. Predictive performance of the four machine learning models for END in patients with acute mild-to-moderate ischemic stroke

Four models (Logistic regression, SVM, random forest, and XGBoost) were constructed based on admission baseline indicators. All achieved AUC values between 0.696 and 0.728, indicating moderate predictive performance. DeLong tests showed no significant differences in AUC among the four models, suggesting comparable overall predictive ability.

The Logistic regression model, as a classical approach, allows intuitive access to regression coefficients for each variable and offers strong interpretability. Its specificity was 0.634. The SVM model demonstrated balanced and stable performance with good generalization under small-sample, high-dimensional data. The random forest model achieved both the highest F1 score (0.469) and the highest specificity (0.687), supporting a comparatively balanced trade-off between sensitivity and specificity. The XGBoost model had the highest sensitivity (0.811), which may be useful when minimizing missed high-risk patients is prioritized, although its lower specificity results in more false-positive classifications. Traditional regression models can be presented as visualized nomograms, whereas ensemble machine learning models have advantages in handling nonlinear features and variable interactions. The negative predictive values of all four models

exceeded 0.90, supporting their potential value for identifying patients at relatively low predicted risk. The absence of statistically significant AUC differences in the DeLong comparisons indicates that these performance differences should be interpreted as model-specific operating characteristics rather than evidence of overall superiority.

4.3. Feature importance based on SHAP algorithm and the predictive value of stress hyperglycemia

The SHAP algorithm was applied to the XGBoost model to quantify feature contributions, addressing the interpretability limitation of black-box models. The top five predictors were ASPECTS, BMI, admission diastolic blood pressure, admission systolic BP, and baseline NIHSS. The stress hyperglycemia ratio ranked seventh and was the primary predictor of interest in this study.

ASPECTS is a well-established quantitative CT-based indicator of early ischemic burden that allows rapid assessment of the extent of early ischemic brain tissue injury. Lower scores indicate larger infarct extent and higher risk of penumbral injury, consistent with its highest predictive weight in this study. Elevated BMI and abnormal admission blood pressure increase END risk through vascular endothelial injury, chronic inflammation, and hemodynamic disturbances, corroborating the baseline differences observed. Baseline NIHSS directly reflects initial deficit severity; greater deficits are associated with higher rates of thrombus extension, cerebral edema, and hemorrhagic transformation.

This study innovatively incorporated the stress hyperglycemia ratio as a predictive feature. The SHAP dependence plot confirmed a significant nonlinear positive correlation with END risk. As SHR increased, oxidative stress, blood-brain barrier disruption, and systemic inflammation progressively intensified, elevating END risk. Compared with a single blood glucose measurement, SHR better reflects the overall post-stroke stress metabolic state and avoids prediction bias from isolated glucose fluctuations. Sex had the lowest predictive contribution, suggesting that after adjustment for confounders, sex was not an independent predictor. The SHAP analysis not only clarified the predictive weights and directions of risk factors but also validated the clinical rationale for incorporating stress hyperglycemia, providing a theoretical basis for future mechanistic studies and model optimization.

4.4. Clinical decision-making value of the two prediction models: insights from decision curve analysis

Decision curve analysis was performed within the 0 to 0.6 risk threshold range to evaluate the net clinical benefit of the Logistic regression and XGBoost models. The clinical utility was threshold-dependent rather than uniform across the entire interval. The Logistic regression curve declined to approximately the zero-net-benefit line at a threshold of about 0.27, whereas the XGBoost curve remained positive over a wider portion of the range and extended to approximately 0.481, consistent with the interval annotated in **Figure 4**. Beyond these points, the corresponding net benefit decreased, so neither model should be interpreted as providing positive net benefit throughout the full 0 to 0.6 range.

These DCA findings suggest that XGBoost may retain clinical net benefit over a broader portion of the evaluated threshold range than Logistic regression, but they do not support a claim of uniformly higher net benefit or superiority across all thresholds. This distinction is consistent with the DeLong analysis, which detected no statistically significant AUC differences among the four models. In practice, model choice should therefore depend on the intended operating point: XGBoost may be useful when sensitivity is prioritized,

whereas the comparatively simple Logistic regression model may remain attractive when interpretability is important. Both models had negative predictive values above 0.90, indicating potential value for identifying patients at relatively low predicted risk. The DCA results support further evaluation of these models as threshold-dependent decision-support tools; prospective external validation is still required before broad clinical implementation.

5. Conclusion

In summary, the risk of END in patients with acute mild-to-moderate ischemic stroke can be predicted using machine learning models constructed from routine admission baseline characteristics. All four models demonstrated moderate discriminative performance, with no statistically significant AUC differences detected by the DeLong test. XGBoost showed the highest sensitivity and retained positive net benefit over a broader portion of the evaluated DCA threshold range, extending to approximately 0.481, whereas this benefit did not persist across the entire 0 to 0.6 interval. ASPECTS showed the highest SHAP importance for END prediction, while the stress hyperglycemia ratio exhibited a nonlinear positive association with the model-predicted risk of END and contributed meaningfully to prediction. These findings support the potential use of machine learning for risk stratification, while external prospective validation and careful consideration of treatment-related confounding are needed before wider clinical application.

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Lin L, Jiang LJ, Wang XG, et al., 2023, Effect of Leukoaraiosis on Intravenous Thrombolysis Efficacy in Acute Mild-to-moderate Cerebral Infarction. *Cardiovascular and Cerebrovascular Disease Prevention and Treatment*, 23(7): 32-36.
- [2] Yang W, Kim KY, Lee JK, et al., 2026, Abstract DP329: Impact of White Matter Hyperintensity Burden on Response to Induced Hypertension Therapy in Small Vessel Occlusion Stroke with Early Neurological Deterioration, *Stroke*, 57(Suppl_1).
- [3] Khasiyev F, Pandit M, Juman S, et al., 2026, Abstract WP118: Exploring Intracranial Arterial Variants and Early Neurological Deterioration in Lacunar Stroke, *Stroke*, 57(Suppl_1).
- [4] Wang N, Lin H, Deng B, et al., 2026, Comparative the Predictive Value of Neutrophil-to-HDL Ratio and Monocyte-to-HDL Ratio for Early Neurological Deterioration in Acute Ischemic Stroke Patients Following Intravenous Thrombolysis. *Neurological Sciences*, 47(5):431.
- [5] Zhong X, Yang J, Xu R, et al., 2026, Anticoagulant or Antithrombotic Therapy after Early Neurological Deterioration in Patients with Branch Atheromatous Disease Receiving Dual Antiplatelet Therapy, *International Journal of Clinical Practice*, 8636214, 7 pages.
- [6] Zhong X, Chen Z, Cao T, et al., 2026, Development and Validation of a Nomogram to Predict Early Neurological Deterioration in Patients with Acute Ischemic Stroke and Type 2 Diabetes Mellitus: the Pivotal Role of Glycemic Variability and Thrombo-inflammation, *Neurological Sciences*, 47(6):544.

- [7] Font MN, García-Sánchez SM, Catalan Stroke Code and Reperfusion Consortium, 2026, Impact of Arterial Occlusion Location on Early Neurological Deterioration in Patients with Minor Stroke Treated with Intravenous Thrombolysis, The MINORCAT-END Study. *Neurological Sciences*, 47(1): 12.
- [8] Xu S, Jiang Z, Zhou Y, et al., 2026, The Effect of Indobufen and Aspirin on Platelet Function in Acute Ischemic Stroke Patients with Early Neurological Deterioration, *European Journal of Pharmacology*, 1011: 178439.
- [9] Neurology Branch and Cerebrovascular Diseases Group of the Chinese Medical Association, 2026, Chinese Guidelines for the Diagnosis and Treatment of Acute Ischemic Stroke 2023: Stroke Recognition, Management, and Transfer. *Practical Journal of Cardiac Cerebral Pneumal and Vascular Disease*, 34(4): 24.
- [10] Writing Group of the Clinical Guidelines for the Prevention and Treatment of Type 2 Diabetes Mellitus in the Elderly in China, Geriatric Endocrinology and Metabolism Branch of the Chinese Geriatrics Society, Geriatric Endocrinology and Metabolism Branch of the Chinese Association of Geriatric Research, et al., 2026, Clinical Guidelines for the Prevention and Treatment of Type 2 Diabetes Mellitus in the Elderly in China (2026 edition) Released in Beijing, *Diabetes Friend*, (5): 7.
- [11] Huang S, Zou DJ., 2025, Treatment of Type 2 Diabetes Mellitus: Should Metformin Step Down to Second-line? Reflections on the Glucose-lowering Pathway of the Chinese Guidelines for Diabetes Prevention and Treatment (2024 edition). *Chinese Journal of Diabetes Mellitus*, 17(10): 1261-1266.
- [12] Zhou XY, 2025, Dietary Habits Survey of Patients with Type 2 Diabetes Mellitus Complicated with Cardiovascular Disease: Based on the Chinese Guidelines for Prevention and Treatment of Type 2 Diabetes (2020 edition), Xi'an Medical University.
- [13] Khasiyev F, Pandit M, Jumani S, et al., 2026, Abstract DP320: Elevated Neutrophil-to-lymphocyte Ratio Does Not Predict Early Neurological Deterioration in Patients with Acute Lacunar Infarct. *Stroke*, 57(Suppl_1).
- [14] Stutz A, Borei K, Pan A, et al., 2026, Abstract DP382: Increased Blood Pressure Variability over the First 24 Hours of Admission is Independently Associated with Neurological Deterioration among Patients with Acute Ischemic Stroke. *Stroke*, 57(Suppl_1).

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.