



Development and validation of a random forest model to predict functional outcome in patients with intracerebral hemorrhage

Daiquan Gao¹ · Wenliang Feng² · Yuanyuan Qiao³ · Xuebin Jiang⁴ · Yunzhou Zhang¹

Received: 7 November 2022 / Accepted: 20 April 2023 / Published online: 10 May 2023
© Fondazione Società Italiana di Neurologia 2023

Abstract

Objective To develop and validate a machine learning (ML)–based model to predict functional outcome in Chinese patients with intracerebral hemorrhage (ICH).

Methods This retrospective cohort study enrolled patients with ICH between November 2017 and November 2020. The follow-up period ended in February 2021. The study population was divided into training and testing sets with a ratio of 7:3. All variables were included in the least absolute shrinkage and selection operator (LASSO) regression for feature selection. The selected variables were incorporated into the random forest algorithm to construct the prediction model. The predictive performance of the model was evaluated via the area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and calibration curve.

Results A total of 412 ICH patients were included, with 288 in the training set, and 124 in the testing set. Twelve attributes were selected: neurological deterioration, Glasgow Coma Scale (GCS) score at 24 h, baseline GCS score, time from onset to the emergency room, blood glucose, diastolic blood pressure (DBP) change in 24 h, hematoma volume change in 24 h, systemic immune-inflammatory index (SII), systolic blood pressure (SBP) change in 24 h, serum creatinine, serum sodium, and age. In the testing set, the accuracy, AUC, sensitivity, specificity, PPV, and NPV of the model were 0.895, 0.964, 0.872, 0.906, 0.810, and 0.939, respectively. The calibration curves showed a good calibration capability of the model.

Conclusion This developed random forest model performed well in predicting 3-month poor functional outcome for Chinese ICH patients.

Keywords Random forest model · Functional outcome · Intracerebral hemorrhage · Prediction

Introduction

Intracerebral hemorrhage (ICH) is an important public health problem, accounting for about 10–15% of all strokes and leading to high mortality and disability rates [1, 2]. The 1-month fatality of ICH was approximately 40%, and it increases with age [3]. The risk of death in patients is usually estimated based on the clinical experience of attending physicians. Nevertheless, overestimation can cause unnecessary withdrawal or care limitation, the so-called self-fulfilling prophecy, while underestimation can result in prolonged, unnecessary, and expensive treatment [4]. Thus, predictive tools for prognosis of ICH patients are of great importance to get favorable outcomes.

Several prognostic tools have been proposed to predict mortality and functional outcome in ICH, but there are many methodological and reporting deficiencies [5]. Recently, owing to the widespread use of machine learning

Daiquan Gao and Wenliang Feng contributed equally to this work.

✉ Yunzhou Zhang
yuz6578@outlook.com

- ¹ Department of Neurology, Xuanwu Hospital, Capital Medical University, No. 45 Changchun Street, Xicheng District, Beijing 100053, China
- ² Department of Critical Care Medicine, Beijing Fengtai You'anmen Hospital, Beijing 100063, China
- ³ Intensive Care Unit, Affiliated Hospital of Jining Medical University, Jining 272029, Shandong, China
- ⁴ Department of Critical Care Medicine, Renhe Hospital, Beijing 102600, China

(ML) algorithms and increasing computing power, there is a growing interest in ML-enhanced diagnostic and prognostic vision in the medical field. Wang et al. developed an ML-based outcome prediction model by combining initial clinical presentation, laboratory data, and imaging findings [6]. Guo et al. also collected initial information on admission to predict the functional outcome of ICH patients in China [7]. Besides, clinical predictors known to affect ICH outcomes were excluded in predicting prognosis for ICH [8]. For instance, ICH can induce the inflammatory response [9], and inflammation is associated with poor clinical outcomes in patients with ICH [10, 11]. Neurological deterioration affects approximately one-third of patients with spontaneous ICH and increases the risk of death and dependency [12, 13]. These factors could provide potential benefits for the prognostic prediction of ICH, indicating that more accurate models could be established by including more variables. Moreover, no widely recognized tool is available in predicting the outcome of Chinese ICH patients [14]. To our knowledge, no prognostic model for ICH applying ML is constructed considering inflammation and changes in patients' condition during treatment. This study aimed to develop and validate an ML-based model to predict functional outcome in Chinese patients with ICH in view of patients' basic characteristics, laboratory tests, imaging findings, and condition changes in treatment.

Methods

Study population

This retrospective cohort study enrolled patients with ICH from two hospitals between November 2017 and November 2020. The follow-up period ended in February 2021. The institutional review boards of the two hospitals approved the study (ethics approval number: [2019]085 and 2021C023). Oral informed consent was obtained from the patients or their guardians. The study was performed in accordance with the Declaration of Helsinki. Inclusion criteria included the following: (1) the diagnosis of ICH complied with the American Heart Association/American Stroke Association ICH guidelines published in 2010; (2) the age of patients was ≥ 18 years; (3) the medical records of patients were complete (including baseline data, laboratory tests, imaging data, treatment records and prognostic data). Exclusion criteria were as follows: (1) patients had cerebral hemorrhage caused by traumatic brain injury, brain tumor, or cerebrovascular malformation; (2) patients took anticoagulant or antiplatelet drugs before the onset; (3) patients died within 7 days of hospitalization; (4) patients had congenital or acquired coagulation factor deficiencies or platelet abnormalities.

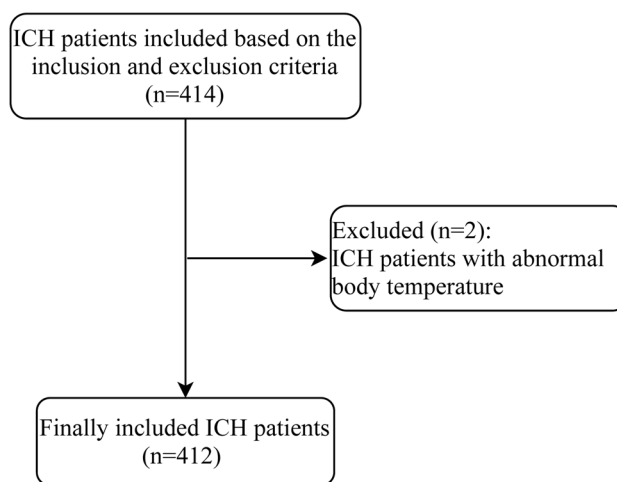


Fig. 1 Flow chart of study population selection. ICH, intracerebral hemorrhage

Data collection

The outcome was a poor functional outcome defined by the Modified Rankin Scale (mRS) score [15] of 3–6 at 3 months, which was collected via follow-up phone interviews. The mRS score of 0–2 was defined as a good functional outcome. We also collected demographic data (gender, age, height, weight, body mass index (BMI), medical history (hypertension, diabetes, hyperlipidemia, atrial fibrillation, acute coronary events, other diseases), smoking history, alcohol use history); baseline (on admission) vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), body temperature, Glasgow Coma Scale (GCS) score); baseline disease characteristics (time from onset to emergency room, hematoma volume, hemorrhage site, intraventricular hemorrhage, intraventricular hemorrhage extension, subarachnoid extension, blend sign, spot sign, leukodystrophy, cerebral lacuna, brain atrophy); baseline laboratory test data (white blood cell count, platelet count, lymphocyte count, neutrophil count, serum potassium, serum sodium, serum calcium, hemoglobin, blood glucose, serum creatinine, activated partial thromboplastin time (APTT), international normalized ratio (INR), fibrin, troponin, total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglyceride); inflammatory marker (systemic immune-inflammatory index (SII)); 24-h admission indicators (SBP at 24 h, DBP at 24 h, hematoma volume at 24 h, GCS score at 24 h); treatment-related indicators (SBP change in 24 h, DBP change in 24 h, hematoma volume change in 24 h, GCS score change in 24 h, neurological deterioration (within 7 day), and treatment (medical treatment, surgical treatment)); mortality. SII = platelet

Table 1 Characteristics of study population

Variables	Total (n=412)	Good functional outcome group (n=273)	Poor functional outcome group (n=139)	Statistics	P
Gender, n (%)				$\chi^2=0.318$	0.573
Male	268 (65.05)	175 (64.10)	93 (66.91)		
Female	144 (34.95)	98 (35.90)	46 (33.09)		
Age, years, mean \pm SD	59.42 \pm 12.59	58.92 \pm 13.07	60.41 \pm 11.57	$t=-1.13$	0.257
Height, cm, mean \pm SD	167.29 \pm 7.81	167.23 \pm 7.97	167.42 \pm 7.50	$t=-0.24$	0.809
Weight, kg, mean \pm SD	71.43 \pm 14.14	71.71 \pm 13.81	70.87 \pm 14.80	$t=0.57$	0.567
BMI, mean \pm SD	25.42 \pm 4.13	25.55 \pm 4.00	25.16 \pm 4.36	$t=0.91$	0.366
Hypertension, n (%)				$\chi^2=1.013$	0.314
No	138 (33.50)	96 (35.16)	42 (30.22)		
Yes	274 (66.50)	177 (64.84)	97 (69.78)		
Diabetes, n (%)				$\chi^2=0.404$	0.525
No	343 (83.25)	225 (82.42)	118 (84.89)		
Yes	69 (16.75)	48 (17.58)	21 (15.11)		
Hyperlipidemia, n (%)				$\chi^2=0.202$	0.653
No	382 (92.72)	252 (92.31)	130 (93.53)		
Yes	30 (7.28)	21 (7.69)	9 (6.47)		
Atrial fibrillation, n (%)				-	1.000
No	408 (99.03)	270 (98.90)	138 (99.28)		
Yes	4 (0.97)	3 (1.10)	1 (0.72)		
Acute coronary events, n (%)				-	1.000
No	408 (99.03)	270 (98.90)	138 (99.28)		
Yes	4 (0.97)	3 (1.10)	1 (0.72)		
Other medical history, n (%)				$\chi^2=6.114$	0.013
No	304 (73.79)	191 (69.96)	113 (81.29)		
Yes	108 (26.21)	82 (30.04)	26 (18.71)		
Smoking history, n (%)				$\chi^2=1.152$	0.283
No	267 (64.81)	172 (63.00)	95 (68.35)		
Yes	145 (35.19)	101 (37.00)	44 (31.65)		
Alcohol use history, n (%)				$\chi^2=1.143$	0.285
No	261 (63.35)	168 (61.54)	93 (66.91)		
Yes	151 (36.65)	105 (38.46)	46 (33.09)		
SBP on admission, mmHg, mean \pm SD	164.21 \pm 29.12	159.66 \pm 26.35	173.16 \pm 32.17	$t=-4.27$	<0.001
DBP on admission, mmHg, mean \pm SD	93.89 \pm 18.18	91.95 \pm 17.17	97.71 \pm 19.52	$t=-3.07$	0.002
Body temperature on admission, °C, mean \pm SD	36.68 \pm 0.61	36.61 \pm 0.43	36.82 \pm 0.83	$t=-2.82$	0.005
GCS score on admission, $M (Q_1, Q_3)$	11.00 (5.00, 15.00)	14.00 (8.00, 15.00)	5.00 (3.00, 11.00)	$Z=-9.544$	<0.001
Time from onset to the emergency room, h, $M (Q_1, Q_3)$	15.50 (3.00, 48.00)	24.00 (5.00, 48.00)	5.00 (2.00, 24.00)	$Z=-6.511$	<0.001
Hematoma volume, mL, $M (Q_1, Q_3)$	20.00 (10.00, 35.00)	16.00 (10.00, 28.00)	30.00 (15.00, 60.00)	$Z=5.289$	<0.001
Hemorrhage site, n (%)				$\chi^2=7.671$	0.104
Basal ganglia	200 (48.54)	127 (46.52)	73 (52.52)		
Brainstem	25 (6.07)	13 (4.76)	12 (8.63)		
Lobe	74 (17.96)	50 (18.32)	24 (17.27)		
Cerebellum	23 (5.58)	14 (5.13)	9 (6.47)		
Other	90 (21.84)	69 (25.27)	21 (15.11)		
Intraventricular hemorrhage, n (%)				$\chi^2=29.733$	<0.001
No	261 (66.24)	194 (75.78)	67 (48.55)		
Yes	133 (33.76)	62 (24.22)	71 (51.45)		

Table 1 (continued)

Variables	Total (n=412)	Good functional outcome group (n=273)	Poor functional outcome group (n=139)	Statistics	P
Intraventricular hemorrhage extension, n (%)				$\chi^2=0.218$	0.640
No	385 (93.45)	254 (93.04)	131 (94.24)		
Yes	27 (6.55)	19 (6.96)	8 (5.76)		
Subarachnoid extension, n (%)				-	0.724
No	403 (97.82)	266 (97.44)	137 (98.56)		
Yes	9 (2.18)	7 (2.56)	2 (1.44)		
Blend sign, n (%)				$\chi^2=5.473$	0.019
No	360 (87.38)	246 (90.11)	114 (82.01)		
Yes	52 (12.62)	27 (9.89)	25 (17.99)		
Spot sign, n (%)				$\chi^2=8.610$	0.003
No	322 (78.16)	225 (82.42)	97 (69.78)		
Yes	90 (21.84)	48 (17.58)	42 (30.22)		
Leukodystrophy, n (%)				$\chi^2=1.682$	0.195
No	307 (74.51)	198 (72.53)	109 (78.42)		
Yes	105 (25.49)	75 (27.47)	30 (21.58)		
Cerebral lacuna, n (%)				$\chi^2=0.059$	0.808
No	273 (66.26)	182 (66.67)	91 (65.47)		
Yes	139 (33.74)	91 (33.33)	48 (34.53)		
Brain atrophy, n (%)				$\chi^2=7.319$	0.007
No	353 (85.68)	243 (89.01)	110 (79.14)		
Yes	59 (14.32)	30 (10.99)	29 (20.86)		
White blood cell count, $10^9/L$, $M (Q_1, Q_3)$	8.80 (7.12, 12.01)	8.36 (6.70, 11.20)	10.05 (7.68, 13.50)	$Z=4.273$	<0.001
Platelet count, $10^9/L$, mean \pm SD	216.44 \pm 63.32	221.29 \pm 63.31	206.94 \pm 62.47	$t=2.19$	0.029
Lymphocyte count, $10^9/L$, $M (Q_1, Q_3)$	1.37 (0.93, 1.83)	1.44 (1.01, 1.85)	1.18 (0.86, 1.71)	$Z=-2.600$	0.009
Neutrophil count, $10^9/L$, $M (Q_1, Q_3)$	7.09 (4.85, 9.61)	6.42 (4.57, 8.89)	8.38 (6.13, 11.58)	$Z=4.985$	<0.001
Serum potassium, mmol/L, $M (Q_1, Q_3)$	3.96 (3.72, 4.21)	3.97 (3.72, 4.20)	3.91 (3.70, 4.25)	$Z=0.182$	0.856
Serum sodium, mmol/L, mean \pm SD	140.24 \pm 5.79	139.32 \pm 4.99	142.06 \pm 6.75	$t=-4.24$	<0.001
Serum calcium, mmol/L, mean \pm SD	2.09 \pm 0.31	2.15 \pm 0.23	1.97 \pm 0.40	$t=5.05$	<0.001
Hemoglobin, g/L, mean \pm SD	136.42 \pm 19.72	137.47 \pm 18.30	134.37 \pm 22.16	$t=1.42$	0.157
Blood glucose, mmol/L, $M (Q_1, Q_3)$	6.40 (5.44, 8.02)	6.19 (5.36, 7.45)	6.80 (5.90, 8.90)	$Z=3.801$	<0.001
Serum creatinine, $\mu\text{mol/L}$, $M (Q_1, Q_3)$	60.05 (51.40, 73.00)	60.00 (51.00, 72.00)	61.90 (52.00, 77.80)	$Z=1.175$	0.240
APTT, mean \pm SD	32.06 \pm 6.67	32.83 \pm 7.13	30.56 \pm 5.38	$t=3.62$	<0.001
INR, $M (Q_1, Q_3)$	1.02 (0.96, 1.09)	1.01 (0.96, 1.07)	1.05 (0.98, 1.11)	$Z=3.773$	<0.001
Fibrin, $M (Q_1, Q_3)$	3.30 (2.70, 4.10)	3.40 (2.70, 4.13)	3.20 (2.60, 3.80)	$Z=-1.713$	0.087
Troponin, $\mu\text{g/L}$, $M (Q_1, Q_3)$	0.12 (0.01, 1.04)	0.84 (0.01, 1.11)	0.02 (0.01, 0.67)	$Z=-4.039$	<0.001
Total cholesterol, mg/dL, mean \pm SD	4.22 \pm 1.15	4.38 \pm 1.22	3.91 \pm 0.93	$t=4.28$	<0.001
Low-density lipoprotein cholesterol, mg/dL, $M (Q_1, Q_3)$	2.52 (1.93, 3.11)	2.64 (2.05, 3.22)	2.34 (1.76, 2.95)	$Z=-3.584$	<0.001
High-density lipoprotein cholesterol, mg/dL, mean \pm SD	1.20 \pm 0.33	1.22 \pm 0.33	1.18 \pm 0.31	$t=1.16$	0.248
Triglyceride, mg/dL, $M (Q_1, Q_3)$	1.17 (0.85, 1.63)	1.24 (0.88, 1.63)	1.06 (0.75, 1.60)	$Z=-2.135$	0.033
SII, $M (Q_1, Q_3)$	1062.02 (655.82, 1894.63)	925.15 (585.26, 1695.42)	1398.10 (785.86, 2407.53)	$Z=4.042$	<0.001
SBP at 24 h, mean \pm SD	112.32 \pm 27.54	107.08 \pm 27.26	122.60 \pm 25.17	$t=-5.60$	<0.001
SBP change in 24 h, $M (Q_1, Q_3)$	51.00 (35.00, 69.00)	50.00 (40.00, 68.00)	52.00 (30.00, 70.00)	$Z=-0.520$	0.603
DBP at 24 h, $M (Q_1, Q_3)$	95.00 (76.00, 140.00)	123.00 (80.00, 149.00)	82.00 (73.00, 114.00)	$Z=-4.683$	<0.001
DBP change in 24 h, $M (Q_1, Q_3)$	40.50 (19.00, 60.00)	45.00 (23.00, 60.00)	30.00 (15.00, 51.00)	$Z=-3.105$	0.002

Table 1 (continued)

Variables	Total (n=412)	Good functional outcome group (n=273)	Poor functional outcome group (n=139)	Statistics	P
Hematoma volume at 24 h, $M (Q_1, Q_3)$	12.00 (7.00, 20.00)	12.00 (7.00, 20.00)	12.00 (7.00, 25.00)	Z=1.215	0.224
Hematoma volume change in 24 h, $M (Q_1, Q_3)$	0.00 (0.00, 15.00)	0.00 (0.00, 10.00)	5.00 (0.00, 44.00)	Z=4.745	<0.001
GCS score at 24 h, $M (Q_1, Q_3)$	11.00 (5.00, 15.00)	14.00 (8.00, 15.00)	5.00 (3.00, 11.00)	Z=-9.686	<0.001
GCS score change in 24 h, $M (Q_1, Q_3)$	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	Z=-0.124	0.902
Neurological deterioration, n (%)				$\chi^2=220.307$	<0.001
No	233 (56.55)	225 (82.42)	8 (5.76)		
Yes	179 (43.45)	48 (17.58)	131 (94.24)		
Treatment, n (%)				$\chi^2=21.462$	<0.001
Medical treatment	262 (63.59)	195 (71.43)	67 (48.20)		
Surgical treatment	150 (36.41)	78 (28.57)	72 (51.80)		
Mortality, n (%)				$\chi^2=97.415$	<0.001
No	360 (87.38)	270 (98.90)	90 (64.75)		
Yes	52 (12.62)	3 (1.10)	49 (35.25)		

–: Fisher's exact test

Poor functional outcome group: participants with the mRS score of 3–6 at 3 months

Good functional outcome group: participants with the mRS score of 0–2 at 3 months

BMI, body mass index; *SBP*, systolic blood pressure; *DBP*, diastolic blood pressure; *GCS*, Glasgow Coma Scale; *APTT*, activated partial thromboplastin time; *INR*, international normalized ratio; *SI*, systemic immune-inflammatory index; *mRS*, Modified Rankin Scale; *SD*, standard deviation

count \times neutrophil count/lymphocyte count. For BP management, the target level to which systolic blood pressure should be lowered in our hospitals is <140 mmHg [16].

Construction and evaluation of the random forest model

The study population was divided into training set and testing set with a ratio of 7:3 using completely random sampling. The training set was used to build the model to predict the functional outcome of ICH patients at 3 months, and the testing set is used for internal validation. All variables were included in the least absolute shrinkage and selection operator (LASSO) regression for variable selection. The selected variables were then incorporated into the random forest algorithm to construct the prediction model. The predictive performance of the random forest model was evaluated via the area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and calibration curve. The 95% confidence intervals (CIs) were estimated.

Statistical analysis

Measurement data with normal distribution were shown as the mean \pm standard deviation (mean \pm SD), and the *t* test was used for intergroup comparisons. Measurement data with skewed distribution were described by the median and quartile ($M (Q_1, Q_3)$), and comparisons of groups were performed using the Wilcoxon rank sum test. Counting data were represented with the number of cases and constituent ratio ($n (%)$), and between-group comparisons were subjected to the chi-square test or Fisher's exact test. Missing values of baseline and 24-h data were imputed using multiple imputation. LASSO regression was applied for multivariate analysis to select variables. The importance ranking of these variables was provided by the random forest model. The receiver operating characteristic (ROC) curve was plotted, and the AUC was calculated. All statistical tests were two-tailed, and $P < 0.05$ was deemed statistically significant. SAS 9.4 (SAS Institute Inc., Cary, NC) was adopted for statistical analysis, and Python 3.8 (Python Software Foundation, Wilmington, DE) was utilized for random forest modeling.

Table 2 Difference between the training and testing sets

Variables	Total (n=412)	Training set (n=288)	Testing set (n=124)	Statistics	P
Gender, n (%)				$\chi^2=0.278$	0.598
Male	268 (65.05)	185 (64.24)	83 (66.94)		
Female	144 (34.95)	103 (35.76)	41 (33.06)		
Age, mean \pm SD	59.42 \pm 12.59	59.20 \pm 12.83	59.95 \pm 12.04	$t=-0.56$	0.578
Height, cm, mean \pm SD	167.29 \pm 7.81	167.21 \pm 7.84	167.48 \pm 7.75	$t=-0.32$	0.746
Weight, kg, mean \pm SD	71.43 \pm 14.14	70.98 \pm 13.94	72.47 \pm 14.59	$t=-0.98$	0.327
BMI, mean \pm SD	25.42 \pm 4.13	25.30 \pm 4.18	25.68 \pm 4.00	$t=-0.87$	0.387
Hypertension, n (%)				$\chi^2=3.772$	0.052
No	138 (33.50)	105 (36.46)	33 (26.61)		
Yes	274 (66.50)	183 (63.54)	91 (73.39)		
Diabetes, n (%)				$\chi^2=1.483$	0.223
No	343 (83.25)	244 (84.72)	99 (79.84)		
Yes	69 (16.75)	44 (15.28)	25 (20.16)		
Hyperlipidemia, n (%)				$\chi^2=0.161$	0.688
No	382 (92.72)	268 (93.06)	114 (91.94)		
Yes	30 (7.28)	20 (6.94)	10 (8.06)		
Atrial fibrillation, n (%)				-	0.587
No	408 (99.03)	286 (99.31)	122 (98.39)		
Yes	4 (0.97)	2 (0.69)	2 (1.61)		
Acute coronary events, n (%)				-	0.587
No	408 (99.03)	286 (99.31)	122 (98.39)		
Yes	4 (0.97)	2 (0.69)	2 (1.61)		
Other medical history, n (%)				$\chi^2=0.733$	0.392
No	304 (73.79)	209 (72.57)	95 (76.61)		
Yes	108 (26.21)	79 (27.43)	29 (23.39)		
Smoking history, n (%)				$\chi^2=0.571$	0.450
No	267 (64.81)	190 (65.97)	77 (62.10)		
Yes	145 (35.19)	98 (34.03)	47 (37.90)		
Alcohol use history, n (%)				$\chi^2=0.015$	0.902
No	261 (63.35)	183 (63.54)	78 (62.90)		
Yes	151 (36.65)	105 (36.46)	46 (37.10)		
SBP on admission, mean \pm SD	164.21 \pm 29.12	162.38 \pm 29.71	168.48 \pm 27.34	$t=-1.96$	0.051
DBP on admission, mean \pm SD	93.89 \pm 18.18	93.32 \pm 18.49	95.20 \pm 17.42	$t=-0.96$	0.337
Body temperature on admission, mean \pm SD	36.68 \pm 0.61	36.68 \pm 0.65	36.67 \pm 0.50	$t=0.28$	0.783
GCS score on admission, $M (Q_1, Q_3)$	11.00 (5.00, 15.00)	11.00 (5.00, 15.00)	12.00 (5.00, 15.00)	$Z=0.443$	0.658
Time from onset to the emergency room, h, $M (Q_1, Q_3)$	15.50 (3.00, 48.00)	14.50 (3.00, 48.00)	17.00 (3.00, 48.00)	$Z=-0.025$	0.980
Hematoma volume, mL, $M (Q_1, Q_3)$	20.00 (10.00, 35.00)	20.00 (10.00, 32.00)	20.00 (12.00, 40.00)	$Z=0.838$	0.402
Hemorrhage site, n (%)				$\chi^2=4.097$	0.393
Basal ganglia	200 (48.54)	146 (50.69)	54 (43.55)		
Brainstem	25 (6.07)	19 (6.60)	6 (4.84)		
Lobe	74 (17.96)	52 (18.06)	22 (17.74)		
Cerebellum	23 (5.58)	14 (4.86)	9 (7.26)		
Other	90 (21.84)	57 (19.79)	33 (26.61)		
Intraventricular hemorrhage, n (%)				$\chi^2=0.002$	0.966
No	261 (66.24)	185 (66.31)	76 (66.09)		
Yes	133 (33.76)	94 (33.69)	39 (33.91)		
Intraventricular hemorrhage extension, n (%)				$\chi^2=0.144$	0.704

Table 2 (continued)

Variables	Total (n=412)	Training set (n=288)	Testing set (n=124)	Statistics	P
No	385 (93.45)	270 (93.75)	115 (92.74)		
Yes	27 (6.55)	18 (6.25)	9 (7.26)		
Subarachnoid extension, n (%)				-	0.730
No	403 (97.82)	281 (97.57)	122 (98.39)		
Yes	9 (2.18)	7 (2.43)	2 (1.61)		
Blend sign, n (%)				$\chi^2=0.191$	0.662
No	360 (87.38)	253 (87.85)	107 (86.29)		
Yes	52 (12.62)	35 (12.15)	17 (13.71)		
Spot sign, n (%)				$\chi^2=0.001$	0.982
No	322 (78.16)	225 (78.13)	97 (78.23)		
Yes	90 (21.84)	63 (21.88)	27 (21.77)		
Leukodystrophy, n (%)				$\chi^2=0.701$	0.402
No	307 (74.51)	218 (75.69)	89 (71.77)		
Yes	105 (25.49)	70 (24.31)	35 (28.23)		
Cerebral lacuna n (%)				$\chi^2=0.517$	0.472
No	273 (66.26)	194 (67.36)	79 (63.71)		
Yes	139 (33.74)	94 (32.64)	45 (36.29)		
Brain atrophy n (%)				$\chi^2=0.715$	0.398
No	353 (85.68)	244 (84.72)	109 (87.90)		
Yes	59 (14.32)	44 (15.28)	15 (12.10)		
White blood cell count, $M(Q_1, Q_3)$	8.80 (7.12, 12.01)	8.67 (6.82, 12.05)	9.08 (7.55, 12.00)	Z=1.388	0.165
Platelet count, mean \pm SD	216.44 \pm 63.32	216.94 \pm 66.53	215.28 \pm 55.36	t=0.26	0.793
Lymphocyte count, $M(Q_1, Q_3)$	1.37 (0.93, 1.83)	1.38 (0.91, 1.76)	1.35 (1.01, 2.02)	Z=1.257	0.209
Neutrophil count, $M(Q_1, Q_3)$	7.09 (4.85, 9.61)	6.84 (4.64, 9.61)	7.39 (5.29, 9.59)	Z=1.500	0.133
Serum potassium, mean \pm SD	4.03 \pm 1.03	4.07 \pm 1.20	3.93 \pm 0.43	t=1.76	0.079
Serum sodium, mean \pm SD	140.24 \pm 5.79	140.07 \pm 5.81	140.64 \pm 5.73	t=-0.92	0.358
Serum calcium, mean \pm SD	2.09 \pm 0.31	2.09 \pm 0.31	2.09 \pm 0.31	t=-0.13	0.900
Hemoglobin, mean \pm SD	136.42 \pm 19.72	136.30 \pm 20.08	136.71 \pm 18.92	t=-0.20	0.845
Blood glucose, $M(Q_1, Q_3)$	6.40 (5.44, 8.02)	6.45 (5.46, 8.20)	6.29 (5.41, 7.61)	Z=-1.388	0.165
Serum creatinine, $M(Q_1, Q_3)$	60.05 (51.40, 73.00)	59.85 (51.00, 73.00)	62.70 (52.00, 73.75)	Z=1.244	0.214
APTT, mean \pm SD	32.06 \pm 6.67	31.55 \pm 5.41	33.26 \pm 8.86	t=-2.00	0.047
INR, $M(Q_1, Q_3)$	1.02 (0.96, 1.09)	1.02 (0.96, 1.08)	1.03 (0.96, 1.10)	Z=1.015	0.310
Fibrin, mean \pm SD	3.49 \pm 1.13	3.47 \pm 1.13	3.52 \pm 1.15	t=-0.44	0.663
Troponin, $M(Q_1, Q_3)$	0.12 (0.01, 1.04)	0.12 (0.01, 1.05)	0.09 (0.01, 1.04)	Z=-0.197	0.844
Total cholesterol, mean \pm SD	4.22 \pm 1.15	4.26 \pm 1.19	4.13 \pm 1.05	t=1.09	0.275
Low-density lipoprotein cholesterol, $M(Q_1, Q_3)$	2.52 (1.93, 3.11)	2.52 (1.99, 3.11)	2.51 (1.87, 3.12)	Z=-0.736	0.462
High-density lipoprotein cholesterol, mean \pm SD	1.20 \pm 0.33	1.21 \pm 0.34	1.20 \pm 0.30	t=0.30	0.767
Triglyceride, $M(Q_1, Q_3)$	1.17 (0.85, 1.63)	1.19 (0.84, 1.61)	1.15 (0.86, 1.73)	Z=0.355	0.723
SII, $M(Q_1, Q_3)$	1062.02 (655.82, 1894.63)	1029.61 (632.38, 1942.42)	1091.20 (684.93, 1805.16)	Z=0.199	0.842
SBP at 24 h, mean \pm SD	112.32 \pm 27.54	112.18 \pm 27.58	112.65 \pm 27.56	t=-0.16	0.874
SBP change in 24 h, $M(Q_1, Q_3)$	51.00 (35.00, 69.00)	50.00 (34.00, 66.50)	57.00 (37.00, 71.00)	Z=1.948	0.051
DBP at 24 h, $M(Q_1, Q_3)$	95.00 (76.00, 140.00)	95.00 (76.00, 140.00)	95.00 (78.50, 150.00)	Z=1.104	0.269
DBP change in 24 h, $M(Q_1, Q_3)$	40.50 (19.00, 60.00)	41.00 (17.00, 57.50)	39.50 (21.50, 61.00)	Z=0.941	0.346
Hematoma volume at 24 h, $M(Q_1, Q_3)$	12.00 (7.00, 20.00)	12.00 (7.00, 20.00)	10.50 (7.00, 19.50)	Z=-0.997	0.319
Hematoma volume change in 24 h, $M(Q_1, Q_3)$	0.00 (0.00, 15.00)	0.00 (0.00, 15.00)	1.50 (0.00, 20.00)	Z=0.500	0.617
GCS score at 24 h, $M(Q_1, Q_3)$	11.00 (5.00, 15.00)	11.00 (5.00, 15.00)	12.00 (5.00, 15.00)	Z=0.648	0.517
GCS score change in 24 h, $M(Q_1, Q_3)$	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	Z=-0.496	0.620

Table 2 (continued)

Variables	Total (n=412)	Training set (n=288)	Testing set (n=124)	Statistics	P
Neurological deterioration, n (%)				$\chi^2=1.620$	0.203
No	233 (56.55)	157 (54.51)	76 (61.29)		
Yes	179 (43.45)	131 (45.49)	48 (38.71)		
Treatment, n (%)				$\chi^2=0.406$	0.524
Medical treatment	262 (63.59)	186 (64.58)	76 (61.29)		
Surgical treatment	150 (36.41)	102 (35.42)	48 (38.71)		

-.: Fisher's exact test

BMI, body mass index; *SBP*, systolic blood pressure; *DBP*, diastolic blood pressure; *GCS*, Glasgow Coma Scale; *APTT*, activated partial thromboplastin time; *INR*, international normalized ratio; *SII*, systemic immune-inflammatory index; *SD*, standard deviation

Results

Characteristics of the study population

A total of 414 patients with ICH were retrospectively reviewed according to the inclusion and exclusion criteria. Two patients with abnormal body temperature (3.6 and 70°C, respectively) were excluded. In the end, 412 ICH patients were included for this study (Figure 1), with 268 (65.05%) males and 144 (34.95%) females. The mean age was 59.42 ± 12.59 years. Most of the patients (66.50%) had hypertension. The median GCS score on admission was 11. The median baseline hematoma volume was 20.00 mL. The majority of patients (63.59%) received medical treatment. The features of ICH patients are presented in Table 1. According to different mRS scores, these patients were divided into the good functional outcome group ($n = 273$) and the poor functional outcome group ($n = 139$). The results of comparisons showed that the two groups were significantly different in other medical history (except history of hypertension,

diabetes, hyperlipidemia, atrial fibrillation, and acute coronary events), SBP, DBP, body temperature, baseline GCS score, time from onset to emergency room, baseline hematoma volume, intraventricular hemorrhage, blend sign, spot sign, brain atrophy, serum sodium, serum calcium, blood glucose, APTT, troponin, total cholesterol, low-density lipoprotein cholesterol, triglyceride, SII, SBP at 24 h, DBP at 24 h, GCS score at 24 h, DBP change in 24 h, hematoma volume change in 24 h, neurological deterioration, treatment, and mortality (all $P < 0.05$). The general mortality of patients was 12.62%; the mortality of the good and poor functional outcome groups was 1.10% and 35.25%, respectively. As shown in Table 2, the training set contained 288 patients, and the testing set included 124 patients. Generally, no significant difference was found in patient characteristics between the two groups.

Ranking of feature importance

After LASSO regression, 12 features are selected (Figure 2). These features were neurological deterioration, GCS score at 24 h, baseline GCS score, time from onset to emergency room, blood glucose, DBP change in 24 h, hematoma volume change in 24 h, SII, SBP change in 24 h, serum creatinine, serum sodium, and age in descending order of importance, according to the feature importance given in the random forest model (Figure 3).

Predictive performance of the random forest model

Table 3 demonstrates the predictive ability of our developed random forest model for functional outcome at 3 months. In the training set, the cut-off value was 0.429, and the model had the accuracy, AUC, sensitivity, specificity, PPV, and NPV of 0.899 (95%CI 0.865–0.934), 0.960 (95%CI 0.939–0.982), 0.920 (95%CI 0.867–0.973), 0.888 (95%CI 0.843–0.933), 0.814 (95%CI 0.742–0.886), and 0.954 (95%CI 0.923–0.985), respectively (Table 3,

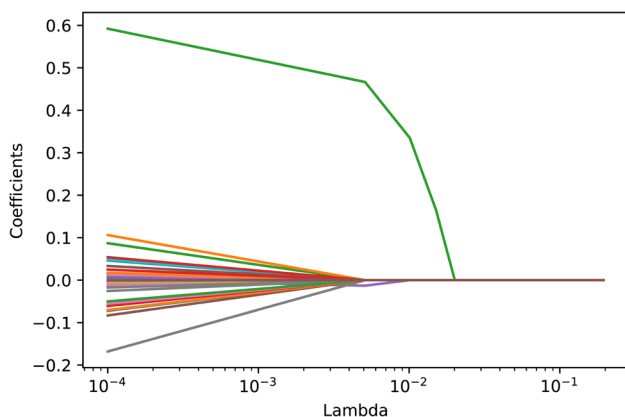


Fig. 2 LASSO feature selection for model construction. LASSO, least absolute shrinkage and selection operator

Fig. 3 Importance ranking of features for model construction. GCS, Glasgow Coma Scale; DBP, diastolic blood pressure; HV, hematoma volume; SII, systemic immune-inflammatory index; SBP, systolic blood pressure

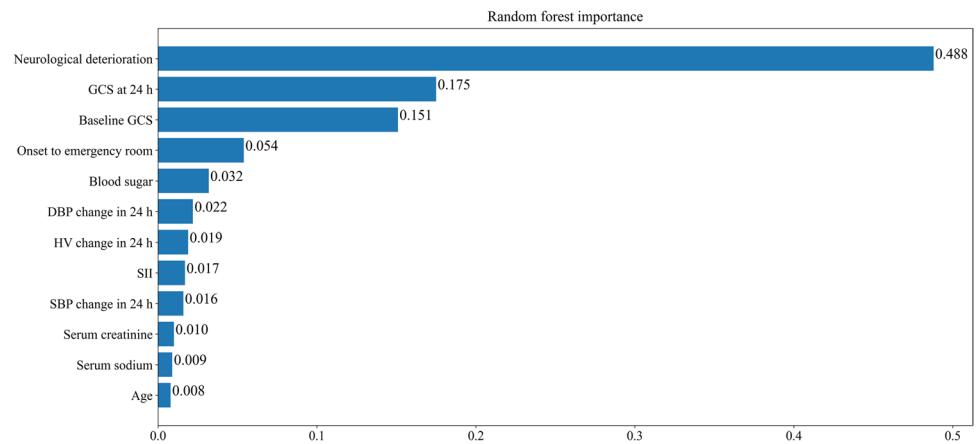


Figure 4). In the testing set for internal validation, the cut-off value was 0.429, and the accuracy, AUC, sensitivity, specificity, PPV, and NPV of the model were 0.895 (95%CI 0.841–0.949), 0.964 (95%CI 0.931–0.998), 0.872 (95%CI 0.767–0.977), 0.906 (95%CI 0.844–0.968), 0.810 (95%CI 0.691–0.928), and 0.939 (95%CI 0.887–0.991), separately (Table 3, Figure 5). These results suggested that this random forest model had a good predictive performance. The calibration curves in the training and testing sets showed a good calibration capability of the model (Figures 6 and 7).

Discussion

Although many prediction tools have been proposed, ML-based prognostic models that take inflammation and changes in patients' condition during treatment into account simultaneously are lacking in predicting poor functional outcome at 3 months for ICH patients. This study developed a random forest model by considering patients' basic features, laboratory test indicators, imaging findings, and condition changes in treatment, with the accuracy, AUC, sensitivity, specificity, PPV, and NPV of 0.895, 0.964, 0.872, 0.906, 0.810, and 0.939 in the validation cohort, respectively, indicating a good predictive capacity of this model in the prediction of 3-month poor functional outcome.

Guo et al. [7] constructed 6 ML-based models to predict the poor functional outcome at 3 months of ICH patients considering the level of monocytes, and proposed that logistic regression (LR) and logistic regression cross-validation (LRCV) models exhibited better predictive performance for functional outcome (AUC = 0.890 and 0.887, respectively) than the ICH score. In this study, we also included the monocyte level to develop a random forest model, with additional consideration of other laboratory variables, such

as blood glucose and serum sodium. Additionally, in the study of Gupta et al. [17], Intracerebral Hemorrhage Outcomes Project 3 (ICHOP₃) was developed for 3-month functional outcome using GCS, National Institutes of Health Stroke Scale, Acute Physiology and Chronic Health Evaluation II, pre-morbid mRS, and hematoma volume to predict the 3-month functional outcome, and had an AUC in the derivation and validation cohorts of 0.89 and 0.745, separately. A prior study by Hall et al. [18] incorporated hematoma volume at hospital admission, hematoma expansion, intraventricular hemorrhage, overall ICH score, and GCS for prediction models of unfavorable functional outcome at 3 months, exhibiting the AUC of 0.75 for decision tree and 0.82 for random forest. The age and neurologic deficit (PLAN) score based on baseline information, as reported by Du et al. [19], had an AUC of 0.84 when predicting functional outcome (mRS 5–6) at discharge. The developed ML-based model herein displayed good discrimination and calibration, with the AUC of 0.964, considering basic characteristics, laboratory results, imaging findings, and condition changes in treatment of ICH patients. Qiu et al. [20] demonstrated that high serum S100A12 levels at admission, which are highly associated with the extent of inflammatory response, severity, and early neurologic deterioration in ICH patients, significantly discriminated a poor functional outcome, with the AUC of 0.794. Copeptin and location-specific differences in hematoma volume were also reported to have prognostic value for the 90-day functional outcome, and great discrimination of functional dependence was found by incorporating one of these two into a prediction model [21, 22]. Future work can consider incorporating serum S100A12 levels, plasma copeptin levels, and interactions between hematoma volume and location to optimize prediction models.

Table 3 Predictive performance of the random forest model

	Cut-off	Accuracy (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)
Training set (n=288)	0.429	0.899 (0.865–0.934)	0.960 (0.939–0.982)	0.920 (0.867–0.973)	0.888 (0.843–0.933)	0.814 (0.742–0.886)	0.954 (0.923–0.985)
Testing set (n=124)	0.429	0.895 (0.841–0.949)	0.964 (0.931–0.998)	0.872 (0.767–0.977)	0.906 (0.844–0.968)	0.810 (0.691–0.928)	0.939 (0.887–0.991)

AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval

According to the feature importance analysis of this random forest model, neurological deterioration was illuminated to significantly contribute to the prediction of 3-month poor functional outcome for ICH patients. Neurological deterioration influences about one-third of patients with ICH and elevates the risk of death and dependence [13, 23, 24]. Law et al. [12] identified that neurological deterioration was independently associated with an mRS of > 3 and unfavorable functional outcome at day 90. Patients without neurological deterioration showed minimal change in hematoma volume at 24 h (1.5 mL), whereas patients with neurological deterioration had a significantly greater absolute increase in hematoma volume (16 mL) [12], which could worsen functional outcome of patients with ICH. High blood glucose levels were shown to correlate with poor functional outcome in ICH [25]. Hematoma with high blood glucose was found to cause nerve damage and decreased autophagy in ICH

models [26]. Hyperglycemia can increase superoxide production in ICH induced by tissue plasminogen activator [27]. Hyperglycemia can also increase hematoma expansion in ICH models by plasma kallikrein [28]. These may contribute to the unfavorable functional outcome of ICH patients. Glycemic control can improve cerebral hemodynamics and neurochemistry in the perihematomal area in primary basal ganglia hemorrhage [29]. Stroke, including intracerebral hemorrhage, results in neuronal cell death and the release of factors such as damage-associated molecular patterns (DAMPs) that elicit localized inflammation in the injured brain region. Such focal brain inflammation aggravates secondary brain injury by exacerbating blood-brain barrier damage, microvascular failure, brain edema, oxidative stress, and by directly inducing neuronal cell death. In addition to inflammation localized to the injured brain region, a growing body of research suggests that inflammatory responses after

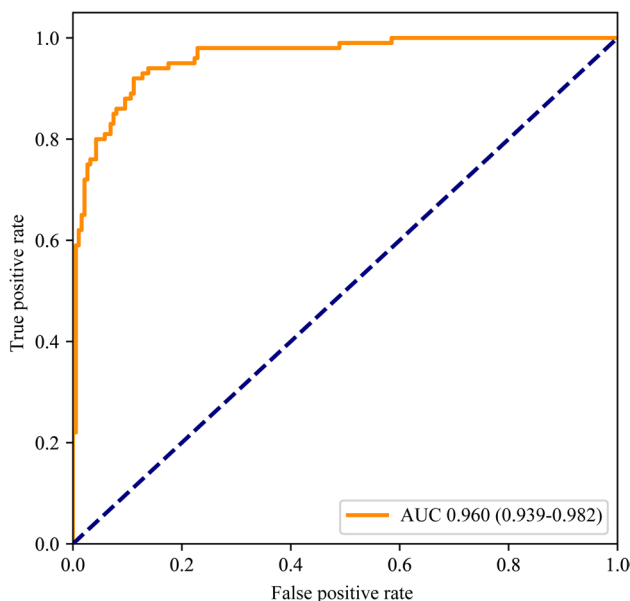


Fig. 4 ROC curve of the random forest model for 3-month functional outcome in the training set. ROC, receiver operating characteristic; AUC, area under the curve

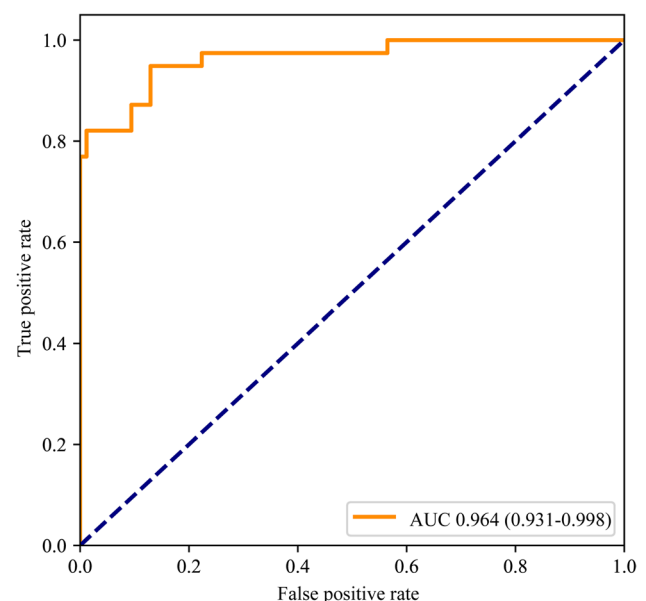


Fig. 5 ROC curve of the random forest model for 3-month functional outcome in the testing set. ROC, receiver operating characteristic; AUC, area under the curve

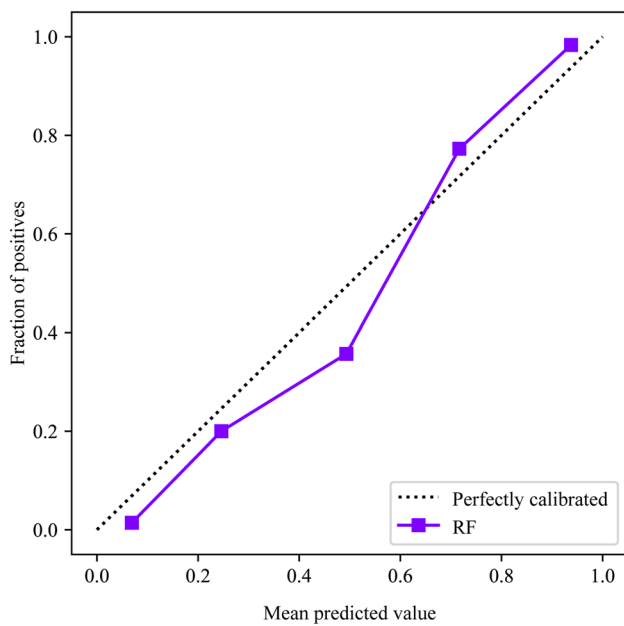


Fig. 6 Calibration curve of the random forest model for 3-month functional outcome in the training set. RF, random forest

a stroke occur and persist throughout the entire brain. Global brain inflammation might continuously shape the evolving pathology after a stroke and affect the patients' long-term neurological outcome [30]. The inflammation and coagulation response after ICH would accelerate the formation of brain edema around hematoma, resulting in a more severe and durable

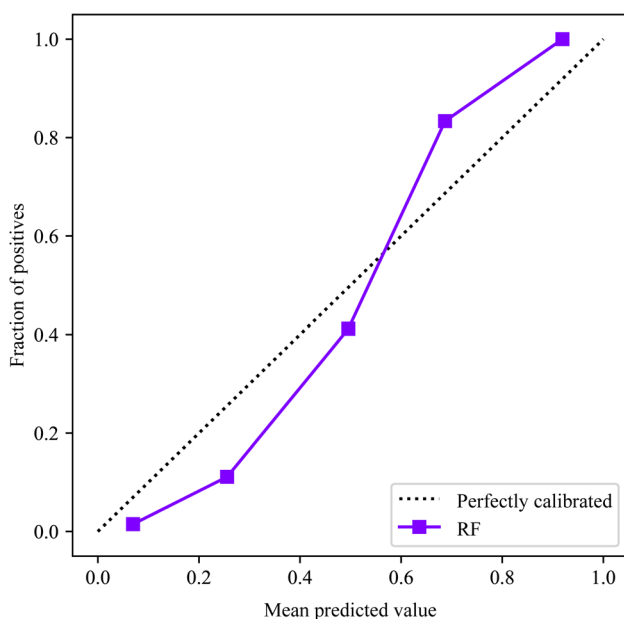


Fig. 7 Calibration curve of the random forest model for 3-month functional outcome in the testing set. RF, random forest

injury [2]. Recent evidence indicates that neuroinflammation plays a critical role in injury expansion and brain damage [31]. As an inflammatory index, SII was highly associated with 90-day functional outcome in patients with ICH and could be used to predict outcomes [32]. Early SII index was reported as an independent predictor of poor outcome at time of hospital discharge for patients with supratentorial spontaneous ICH [10]. With additional attention to these factors, the random forest model performed well in predicting the functional outcome of ICH patients at 3 months.

A random forest model with good predictive capacity was established in the current study, incorporating inflammation and changes in patients' condition during treatment, which may help to identify ICH patients with poor 3-month functional outcome and offer early interventions in clinical practice. There were several limitations. First, external validation is lacking in this paper, which may restrict the generalizability of our findings. Second, the mRS score was collected via follow-up phone interviews. Misjudgment of patients' functional outcome evaluation may be present because interviewees might not be able to give accurate answers during the follow-up phone interview. Although prior studies have shown that assessment of the mRS through the structured telephone is reliable and comparable with face-to-face interviews [33], long-time functional outcome was not assessed in our study. Third, some variables such as neuroimaging and genetic variables, and medication of patients after discharge were not taken into account. Besides, due to the retrospective nature of this study, we could not obtain complete information about specific therapeutic procedures, the department where patients were admitted, the length of hospital stay, and the rehabilitation administered. More studies are needed in predicting functional outcome by considering these variables.

Conclusion

A random forest model was developed and validated to have a good predictive performance for poor functional outcome at 3 months among Chinese ICH patients, which might provide additional assistance for clinicians to predict 3-month functional outcome and offer appropriate medical care. Further studies are required to evaluate the applicability of this model in other populations.

Authors' contributions YZ designed the study. YQ, WF, and XJ collected and analyzed the data. DG and YQ wrote the manuscript. YZ reviewed and edited the manuscript. All the authors read and approved the final manuscript.

Funding This work was supported by the STI2030-Major Projects+2021ZD0204300.

Declarations

Ethical approval The institutional review boards of two hospitals approved the study (ethics approval number: [2019]085 and 2021C023).

Informed consent Informed consent was obtained from all participants in the study.

Conflict of interest None.

References

- Qureshi AI, Mendelow AD, Hanley DF (2009) Intracerebral haemorrhage. *Lancet* (London, England) 373(9675):1632–1644. [https://doi.org/10.1016/s0140-6736\(09\)60371-8](https://doi.org/10.1016/s0140-6736(09)60371-8)
- Zheng H, Chen C, Zhang J, Hu Z (2016) Mechanism and therapy of brain edema after intracerebral hemorrhage. *Cerebrovasc Dis* 42(3–4):155–169. <https://doi.org/10.1159/000445170>
- van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ (2010) Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol* 9(2):167–176. [https://doi.org/10.1016/s1474-4422\(09\)70340-0](https://doi.org/10.1016/s1474-4422(09)70340-0)
- Satopää J, Mustanoja S, Meretoja A, Putaala J, Kaste M, Niemelä M, Tatlisumak T, Strbian D (2017) Comparison of all 19 published prognostic scores for intracerebral hemorrhage. *J Neurol Sci* 379:103–108. <https://doi.org/10.1016/j.jns.2017.05.034>
- Gregório T, Pipa S, Cavaleiro P, Atanásio G, Albuquerque I, Chaves PC, Azevedo L (2018) Prognostic models for intracerebral hemorrhage: systematic review and meta-analysis. *BMC Med Res Methodol* 18(1):145. <https://doi.org/10.1186/s12874-018-0613-8>
- Wang HL, Hsu WY, Lee MH, Weng HH, Chang SW, Yang JT, Tsai YH (2019) Automatic machine-learning-based outcome prediction in patients with primary intracerebral hemorrhage. *Front Neurol* 10:910. <https://doi.org/10.3389/fneur.2019.00910>
- Guo R, Zhang R, Liu R, Liu Y, Li H, Ma L, He M, You C, Tian R (2022) Machine learning-based approaches for prediction of patients' functional outcome and mortality after spontaneous intracerebral hemorrhage. *J Pers Med* 12(1). <https://doi.org/10.3390/jpm12010112>
- Rost NS, Smith EE, Chang Y, Snider RW, Chanderraj R, Schwab K, FitzMaurice E, Wendell L, Goldstein JN, Greenberg SM, Rosand J (2008) Prediction of functional outcome in patients with primary intracerebral hemorrhage: the FUNC score. *Stroke* 39(8):2304–2309. <https://doi.org/10.1161/strokeaha.107.512202>
- Askenase MH, Sansing LH (2016) Stages of the inflammatory response in pathology and tissue repair after intracerebral hemorrhage. *Semin Neurol* 36(3):288–297. <https://doi.org/10.1055/s-0036-1582132>
- Trifan G, Testai FD (2020) Systemic Immune-Inflammation (SII) index predicts poor outcome after spontaneous supratentorial intracerebral hemorrhage. *J Stroke Cerebrovasc Dis* 29(9):105057. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105057>
- Fonseca S, Costa F, Seabra M, Dias R, Soares A, Dias C, Azevedo E, Castro P (2021) Systemic inflammation status at admission affects the outcome of intracerebral hemorrhage by increasing perihematomal edema but not the hematoma growth. *Acta Neurol Belg* 121(3):649–659. <https://doi.org/10.1007/s13760-019-01269-2>
- Law ZK, Dineen R, England TJ, Cala L, Mistri AK, Appleton JP, Ozturk S, Bereczki D, Ciccone A, Bath PM, Sprigg N (2021) Predictors and outcomes of neurological deterioration in intracerebral hemorrhage: results from the TICH-2 randomized controlled trial. *Transl Stroke Res* 12(2):275–283. <https://doi.org/10.1007/s12975-020-00845-6>
- Lord AS, Gilmore E, Choi HA, Mayer SA (2015) Time course and predictors of neurological deterioration after intracerebral hemorrhage. *Stroke* 46(3):647–652. <https://doi.org/10.1161/strokeaha.114.007704>
- Wu S, Wu B, Liu M, Chen Z, Wang W, Anderson CS, Sandercock P, Wang Y, Huang Y, Cui L, Pu C, Jia J, Zhang T, Liu X, Zhang S, Xie P, Fan D, Ji X, Wong KL, Wang L (2019) Stroke in China: advances and challenges in epidemiology, prevention, and management. *Lancet Neurol* 18(4):394–405. [https://doi.org/10.1016/s1474-4422\(18\)30500-3](https://doi.org/10.1016/s1474-4422(18)30500-3)
- Broderick JP, Adeoye O, Elm J (2017) Evolution of the modified rankin scale and its use in future stroke trials. *Stroke* 48(7):2007–2012. <https://doi.org/10.1161/strokeaha.117.017866>
- Anderson CS, Heeley E, Huang Y, Wang J, Stapf C, Delcourt C, Lindley R, Robinson T, Lavados P, Neal B, Hata J, Arima H, Parsons M, Li Y, Wang J, Heritier S, Li Q, Woodward M, Simes RJ et al (2013) Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage. *N Engl J Med* 368(25):2355–2365. <https://doi.org/10.1056/NEJMoa1214609>
- Gupta VP, Garton ALA, Sisti JA, Christophe BR, Lord AS, Lewis AK, Frey HP, Claassen J, Connolly ES Jr (2017) Prognosticating functional outcome after intracerebral hemorrhage: the ICHOP score. *World Neurosurg* 101:577–583. <https://doi.org/10.1016/j.wneu.2017.02.082>
- Hall AN, Weaver B, Liotta E, Maas MB, Faigle R, Mroczek DK, Naidech AM (2021) Identifying modifiable predictors of patient outcomes after intracerebral hemorrhage with machine learning. *Neurocrit Care* 34(1):73–84. <https://doi.org/10.1007/s12028-020-00982-8>
- Du W, Zhao X, Wang Y, Zhang G, Fang J, Pan Y, Liu L, Dong K, Liu G, Wang Y; China National Stroke Registry (CNSR) Investigators (2020) The PLAN score can predict poor outcomes of intracerebral hemorrhage. *Ann Transl Med* 8 (1):14. <https://doi.org/10.21037/atm.2019.11.88>
- Qiu SZ, Zheng GR, Ma CY, Chen B, Huang JJ, Huang G, Hua H (2021) High serum S100A12 levels predict poor outcome after acute primary intracerebral hemorrhage. *Neuropsychiatr Dis Treat* 17:3245–3253. <https://doi.org/10.2147/ndt.S337041>
- Wei ZJ, Ou YQ, Li X, Li H (2014) The 90-day prognostic value of copeptin in acute intracerebral hemorrhage. *Neurol Sci* 35(11):1673–1679. <https://doi.org/10.1007/s10072-014-1809-2>
- Ironside N, Chen CJ, Dreyer V, Christophe B, Buell TJ, Connolly ES (2020) Location-specific differences in hematoma volume predict outcomes in patients with spontaneous intracerebral hemorrhage. *Int J Stroke* 15(1):90–102. <https://doi.org/10.1177/1747493019830589>
- Sun W, Pan W, Kranz PG, Hailey CE, Williamson RA, Sun W, Laskowitz DT, James ML (2013) Predictors of late neurological deterioration after spontaneous intracerebral hemorrhage. *Neurocrit Care* 19(3):299–305. <https://doi.org/10.1007/s12028-013-9894-2>
- Leira R, Dávalos A, Silva Y, Gil-Peralta A, Tejada J, Garcia M, Castillo J (2004) Early neurologic deterioration in intracerebral hemorrhage: predictors and associated factors. *Neurology* 63(3):461–467. <https://doi.org/10.1212/01.wnl.0000133204.81153.ac>
- Zheng J, Yu Z, Ma L, Guo R, Lin S, You C, Li H (2018) Association between blood glucose and functional outcome in

- intracerebral hemorrhage: a systematic review and meta-analysis. *World Neurosurg* 114:e756–e765. <https://doi.org/10.1016/j.wneu.2018.03.077>
26. Liu RY, Wang JJ, Qiu X, Wu JM (2014) Acute hyperglycemia together with hematoma of high-glucose blood exacerbates neurological injury in a rat model of intracerebral hemorrhage. *Neurosci Bull* 30(1):90–98. <https://doi.org/10.1007/s12264-013-1371-6>
 27. Won SJ, Tang XN, Suh SW, Yenari MA, Swanson RA (2011) Hyperglycemia promotes tissue plasminogen activator-induced hemorrhage by increasing superoxide production. *Ann Neurol* 70(4):583–590. <https://doi.org/10.1002/ana.22538>
 28. Liu J, Gao BB, Clermont AC, Blair P, Chilcote TJ, Sinha S, Flaumenhaft R, Feener EP (2011) Hyperglycemia-induced cerebral hematoma expansion is mediated by plasma kallikrein. *Nat Med* 17(2):206–210. <https://doi.org/10.1038/nm.2295>
 29. Ho CL, Ang CB, Lee KK, Ng IH (2008) Effects of glycaemic control on cerebral neurochemistry in primary intracerebral haemorrhage. *J Clin Neurosci* 15(4):428–433. <https://doi.org/10.1016/j.jocn.2006.08.011>
 30. Shi K, Tian DC, Li ZG, Ducruet AF, Lawton MT, Shi FD (2019) Global brain inflammation in stroke. *Lancet Neurol* 18(11):1058–1066. [https://doi.org/10.1016/s1474-4422\(19\)30078-x](https://doi.org/10.1016/s1474-4422(19)30078-x)
 31. Lucke-Wold BP, Logsdon AF, Manoranjan B, Turner RC, McConnell E, Vates GE, Huber JD, Rosen CL, Simard JM (2016) Aneurysmal subarachnoid hemorrhage and neuroinflammation: a comprehensive review. *Int J Mol Sci* 17(4):497. <https://doi.org/10.3390/ijms17040497>
 32. Li Y, Wen D, Cui W, Chen Y, Zhang F, Yuan M, Xiao H, Li H, Ma L, Hu X, You C (2021) The prognostic value of the acute phase systemic immune-inflammation index in patients with intracerebral hemorrhage. *Front Neurol* 12:628557. <https://doi.org/10.3389/fneur.2021.628557>
 33. Merino JG, Lattimore SU, Warach S (2005) Telephone assessment of stroke outcome is reliable. *Stroke* 36(2):232–233. <https://doi.org/10.1161/01.STR.0000153055.43138.2f>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.