

30-Day Mortality Risk Prediction Model for Intracerebral Hemorrhage Patients Combined with Albumin Infusion Characteristics

Xiangjun Chen

Hubei University of Medicine

Zhigang Huang

cx.j645082633@outlook.com

Peking University Shenzhen Hospital

Research Article

Keywords: Acute intracerebral hemorrhage, Prediction model, Machine learning, SHAP analysis

Posted Date: July 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9664933/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

30-Day Mortality Risk Prediction Model for Intracerebral Hemorrhage Patients Combined with Albumin Infusion Characteristics

Author: Xiangjun Chen¹ Zhigang Hang^{2*}

Unit:¹Department of General Practice, Hubei University of Medicine, Shiyan, China

^{2*}Department of General Practice, Peking University Shenzhen Hospital, Shenzhen, China

Email:cxj645082633@outlook.com

Abstract

background: This paper aims to evaluate the 30-day mortality risk of intracerebral hemorrhage patients combined with albumin infusion characteristics using a machine learning model.

Methods: Data were derived from two non-overlapping datasets from the MIMIC-IV data warehouse. The MIMIC-IV dataset was used for training, and the MIMIC-III dataset served as an external validation set. With feature selection performed using Lasso regression, eight machine learning algorithms were employed to build prediction models, and model performance was evaluated based on three dimensions: discrimination, calibration, and clinical utility. SHapley Additive exPlanations (SHAP) analysis was used for interpretability analysis.

Results: The Logistic Regression (LR) model performed best in predicting 30-day mortality risk in patients with intracerebral hemorrhage and characteristics. SHAP analysis enhanced model interpretability. This method provides clinicians with a tool

for individualized risk assessment, aiding in the optimization of albumin infusion strategies and early interventions, thereby reducing mortality risk. The LR model achieved an area under the receiver operating characteristic curve (AUC) greater than 0.8 in both the training and validation sets (training set: 0.814, 95% CI: 0.802-0.825; Internal validation set: 0.811, 95% CI: 0.785-0.837; External validation set: 0.803, 95% CI: 0.781-0.825). The model's SHAP variable importance, ranked from highest to lowest, was: age, Sequential Organ Failure Assessment (SOFA) score, mannitol, white blood cell count, oxygen saturation, glucose, red cell distribution width, weight, anion gap, continuous renal replacement therapy, sodium, albumin infusion.

Conclusion: The Logistic Regression (LR) model performed best in predicting 30-day mortality risk in patients with intracerebral hemorrhage and albumin infusion characteristics. SHAP analysis provided interpretable results of the machine learning results.

Keywords: Acute intracerebral hemorrhage, Prediction model, Machine learning, SHAP analysis

1.Introduction

Acute intracerebral hemorrhage (ICH), as a severe cerebrovascular disease characterized by acute onset and rapid progress associated with extremely high mortality and disability rates^[1]. Epidemiological data indicate that ICH accounts for 10%–15% of all stroke cases, with a 30-day mortality rate as high as 30%–50%, making it one of the major global public health challenges^[2]. Albumin is a single polypeptide chain folded into three domains consisting of 586 amino acids, with a

molecular weight of approximately 67 kDa. It is the most critical protein in human plasma, with physiological functions such as binding and transporting lipids and lipophilic substances, providing nutritional support, and maintaining plasma colloid osmotic pressure^[3]. Albumin exerts neuroprotective effects on the occurrence and development of ischemic stroke, with mechanisms encompassing antioxidant, anti-inflammatory, and anti-edema functions, maintaining vascular permeability, and regulating metabolism^[4,5]. Although it is widely used in the clinical treatment of ICH, existing diagnostic tools, while convenient to operate, fail to fully incorporate variables relating to infusion. Furthermore, they struggle to capture the complex nonlinear relationships and potential interactions between clinical characteristics, laboratory indicators, and prognosis, resulting in limited accuracy in predicting 30-day mortality risk and difficulty in meeting the needs of personalized diagnosis and treatment. Therefore, timely and accurate assessment of the 30-day mortality risk in ICH patients is of crucial significance in formulating clinical intervention plans and improving patient outcomes. In recent years, machine learning technology has demonstrated unique advantages in the field of predicting the prognosis of severe diseases. Compared with traditional statistics, machine learning algorithms such as gradient boosting decision trees, random forests, and neural networks can efficiently process clinical data, accurately identify complex association patterns among multiple variables, and significantly improve the performance and stability of prediction models^[6]. Combining machine learning methods with clinical indicators related to albumin infusion to construct a prediction model for the 30-day mortality of ICH

patients can provide clinicians with a tool for individualized risk assessment. This helps optimize albumin infusion strategies and early intervention plans, thereby reducing mortality risk. This study also provides new technical support and research ideas for precision medicine in ICH patients, holding important clinical value and social significance for improving patient outcomes and the rational allocation of medical resources.

2.Data Sources and Research Subjects

2.1 research design

This study aims to construct a series of machine learning models to predict the 30-day mortality risk of intracerebral hemorrhage (ICH) patients with combined albumin infusion characteristics. This retrospective study included the MIMIC-III (v1.4) and MIMIC-IV (v3.1) databases, associated with the 30-day mortality risk of ICH patients. Subsequently, a series which cover detailed information of intensive care unit (ICU) patients from 2008-2022 and 2001-2012, respectively (Figure 1). Feature selection was performed using the Boruta algorithm, recursive feature elimination, and Lasso regression to identify features most. As machine learning models were developed and optimized, including Logistic Regression (LR), Support Vector Machine (SVM), Gradient Boosting Machine (GBM), Neural Network, Xgboost, Adaboost, LightGBM, and CatBoost. Grid search was used to select the optimal hyperparameters for each model. During the parameter tuning phase, the model with the highest Area Under the Receiver Operating Characteristic Curve (AUC) was selected as the optimal model. The optimal model was selected on the training set and

validated on both internal and external validation sets. The performance of the prediction models was evaluated using metrics such as AUC, sensitivity, specificity, recall, F1 score, and accuracy. Furthermore, decision curve analysis (DCA) and calibration curves were plotted to demonstrate their practical clinical value.

The SHAP method was employed to generate summary plots to demonstrate the contribution of each feature to the prediction results. The SHAP evaluation of the selected case reveals the impact of features on specific samples, aiding in the understanding of the model's decision-making process.

2.2 Research subject

Inclusion criteria: (1) Age ≥ 18 years; (2) First admission to the ICU; (3) ICU stay ≥ 12 hours; (4) Meet the diagnostic criteria for intracerebral hemorrhage.

Exclusion criteria: (1) Missing key laboratory indicators $> 30\%$; (2) Advanced malignancy, end-stage renal disease, brain death (severely affecting patient prognosis); (3) Abnormal recorded vital signs.

2.3 Data Extraction and Cleaning

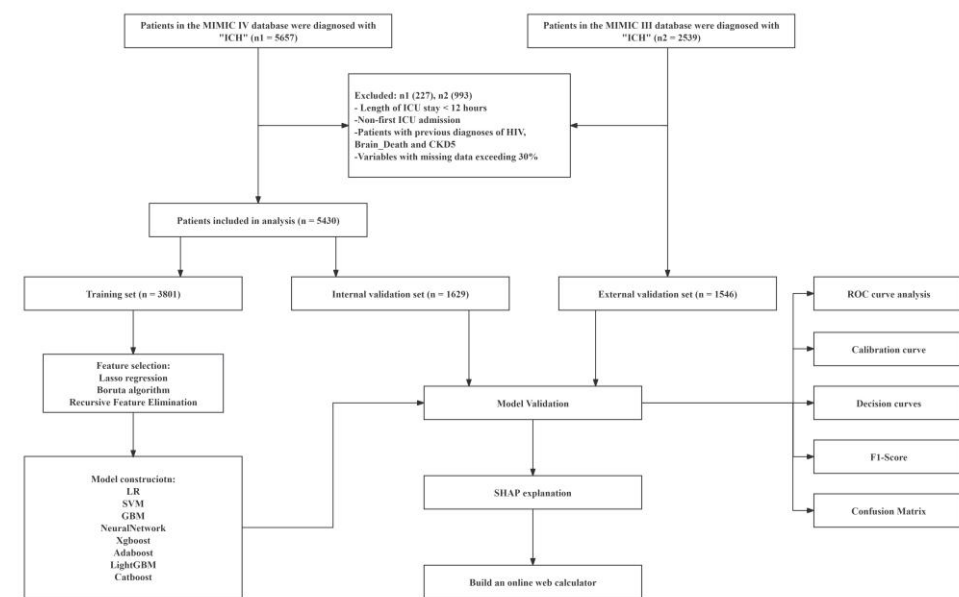
Use PostgreSQL to extract tables such as diagnoses_icd, procedures_icd, labevents, chartevents, prescriptions, admissions, etc. The variables extracted in this study include: (1) Demographic variables: age, gender, race; (2) Clinical scores: Sequential Organ Failure Assessment (SOFA), Acute Physiological And Chronic Health Evaluation III (APACHE III), Simplified Acute Physiology Score II (SAPS II); (3) Blood gas indicators: partial pressure of carbon dioxide ((PCO₂)), partial pressure of

oxygen (PO₂), pH, lactate; (4) Laboratory indicators: International Normalized Ratio (INR), creatinine (Cr), sodium (Na), potassium (K), glucose (Glu); (5) Interventions: albumin infusion, mannitol infusion, mechanical ventilation, continuous renal replacement therapy (CRRT) ; (6) Outcomes: 30-day ICU mortality. Missing value handling: MICE, multiple imputation, or MissForest.

2.4 statistical analysis

Data analysis was performed using R software (version number). Continuous variables were expressed as mean $\pm$ standard deviation or median $\pm$ interquartile range depending on distribution. For normally distributed data, comparisons were made using independent samples t-tests; for non-normally distributed data, the Mann-Whitney U test was used. Categorical variables are presented as frequencies and percentages, and comparisons between groups were performed using the chi-square test or Fisher's exact test.

Figure 1 Flowchart



3.Result

3.1 baseline features

After screening, a total of 5,430 patients from MIMIC-IV were included in the study, of whom 934 died, resulting in a mortality rate of 17.20%. The baseline characteristics of the patients are shown in Table 1. Variables with statistically significant differences between the two groups included: gender ($P < 0.001$), continuous renal replacement therapy ($P < 0.001$), mechanical ventilation ($P < 0.001$), acute kidney injury ($P < 0.001$), hypertension ($P < 0.001$), type 2 diabetes ($P < 0.001$), mannitol ($P < 0.001$), albumin infusion ($P < 0.001$), body weight ($P < 0.001$), age ($P < 0.001$), partial pressure of oxygen ($P < 0.001$), and anion gap ($P < 0.001$), etc.

Table 1 Description of Research Variables

Variables	Total (n = 5430)	Non-death (n = 4496)	death (n = 934)	p
Gender(%)				0.141
yes	2512 (46)	2059 (46)	453 (49)	
no	2918 (54)	2437 (54)	481 (51)	
CRRT (%)				< 0.001
yes	5362 (99)	4465 (99)	897 (96)	
no	68 (1)	31 (1)	37 (4)	
MV(%)				< 0.001
yes	1853 (34)	1609 (36)	244 (26)	
no	3577 (66)	2887 (64)	690 (74)	
Hypertension (%)				0.016
yes	2411 (44)	2030 (45)	381 (41)	
no	3019 (56)	2466 (55)	553 (59)	
AKI(%)				< 0.001
yes	4642 (85)	3924 (87)	718 (77)	
no	788 (15)	572 (13)	216 (23)	
CVA(%)				< 0.001
yes	3888 (72)	3281 (73)	607 (65)	
no	1542 (28)	1215 (27)	327 (35)	
T2DM(%)				< 0.001
yes	4296 (79)	3610 (80)	686 (73)	
no	1134 (21)	886 (20)	248 (27)	
Mannitol(%)				< 0.001
yes	4833 (89)	4144 (92)	689 (74)	
no	597 (11)	352 (8)	245 (26)	
BUN	15 (11, 21)	15 (11, 20)	19 (14, 26)	< 0.001

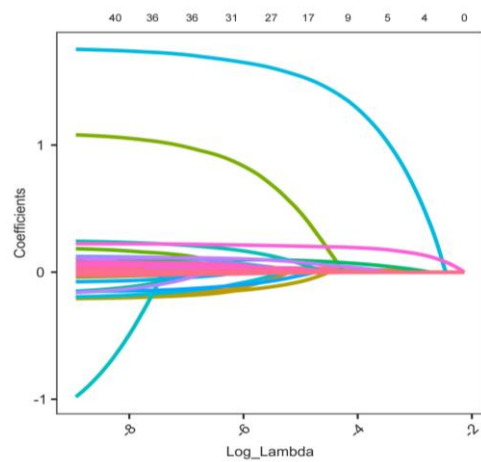
Variables	Total (n = 5430)	Non-death (n = 4496)	death (n = 934)	p
Albumin_Infusio				< 0.001
n(%)				
yes	5047 (93)	4208 (94)	839 (90)	
no	383 (7)	288 (6)	95 (10)	
DBP	73 (63, 84)	73 (63, 84)	71 (59, 83)	< 0.001
MAP	88 (78, 99)	88 (78, 99)	87 (75, 99.75)	0.01
RR	18 (15, 21)	18 (15, 21)	18 (16, 22)	< 0.001
SPO2	98 (96, 100)	98 (96, 100)	99 (96, 100)	< 0.001
Temperature	36.83 (36.56, 37.17)	36.83 (36.56, 37.11)	36.78 (36.44, 37.17)	< 0.001
Hematocrit	36.4 (32.5, 39.9)	36.6 (32.8, 40)	35.4 (30.9, 39.6)	< 0.001
Hemoglobin	12.1 (10.8, 13.4)	12.2 (10.9, 13.5)	11.7 (10.1, 13.17)	< 0.001
RDW	13.7 (13, 14.7)	13.6 (12.9, 14.6)	14.2 (13.3, 15.4)	< 0.001
Calcium	8.7 (8.3, 9.1)	8.7 (8.3, 9.1)	8.6 (8.1, 9)	< 0.001
Chloride	104 (101, 107)	104 (101, 107)	104 (100, 107)	0.987
Weight	75.2 (63.4, 89.1)	75.5 (64, 89.9)	74.05 (60.73, 86.65)	< 0.001
Age	68 (55, 80)	67 (53, 79)	75 (64, 84)	< 0.001
SOFA	3 (2, 4)	2 (1, 4)	4 (2, 7)	< 0.001
GCS	14 (12, 15)	14 (12, 15)	15 (8, 15)	0.724
HR	81 (71, 94)	81 (71, 93)	83 (70, 96)	0.112
Glucose	7.06 (5.89, 8.67)	6.89 (5.83, 8.33)	8.28 (6.74, 10.56)	< 0.001
Potassium	3.9 (3.6, 4.3)	3.9 (3.6, 4.3)	4 (3.6, 4.4)	0.001
Sodium	139 (137, 142)	139 (137, 142)	140 (137, 142)	0.128
Bicarbonate	23 (21, 25)	23 (21, 25)	23 (20, 25)	< 0.001
Magnesium	1.9 (1.7, 2.1)	1.9 (1.7, 2.1)	1.9 (1.7, 2.1)	0.394
Phosphate	3.3 (2.8, 3.8)	3.3 (2.8, 3.8)	3.2 (2.6, 3.9)	0.759
Creatinine	0.8 (0.7, 1.1)	0.8 (0.7, 1)	1 (0.7, 1.3)	< 0.001
INR	1.1 (1.1, 1.3)	1.1 (1.1, 1.2)	1.2 (1.1, 1.4)	< 0.001
PT	12.6 (11.7, 13.9)	12.5 (11.6, 13.7)	13.3 (12.1, 15.4)	< 0.001
APTT	28 (25.6, 31)	27.9 (25.6, 30.7)	28.6 (25.3, 32.2)	< 0.001
Platelet	205 (161, 257)	207 (165, 258.25)	192 (142.25, 249.75)	< 0.001
SBP	132 (118, 147)	132 (119, 146)	132 (114, 149)	0.335
RBC	4.01 (3.55, 4.44)	4.03 (3.58, 4.45)	3.87 (3.36, 4.39)	< 0.001
WBC	10.3 (8, 13.3)	10 (7.8, 12.9)	11.8 (9.03, 15.2)	< 0.001
AG	14 (12, 16)	14 (12, 16)	15 (13, 17)	< 0.001

3.2 Feature selection

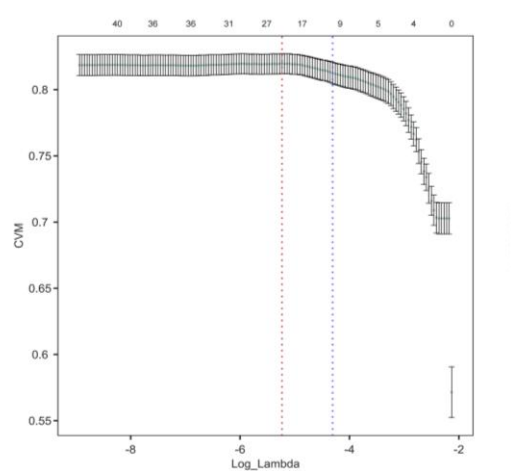
We identified predictive factors for mortality in intracerebral hemorrhage patients using Lasso regression (Figure 2A). By setting the value to one standard deviation

from the minimum λ , where the error remained within one standard error of the minimum, we identified 14 variables with the highest predictive power. Concurrently, we also selected 40 predictive variables using recursive elimination (Figure 2C). Using a Venn diagram, we finally identified 14 predictors that appeared in both feature selection methods: "anion gap", "sodium", "oxygen saturation", "white blood cells", "blood urea nitrogen", "prothrombin time", "age", "continuous renal replacement therapy", "red cell distribution width", "glucose", "Sequential Organ Failure Assessment score", "mannitol", "weight", "albumin infusion" (Figure 2D). These features were ultimately used for model construction.

A



B



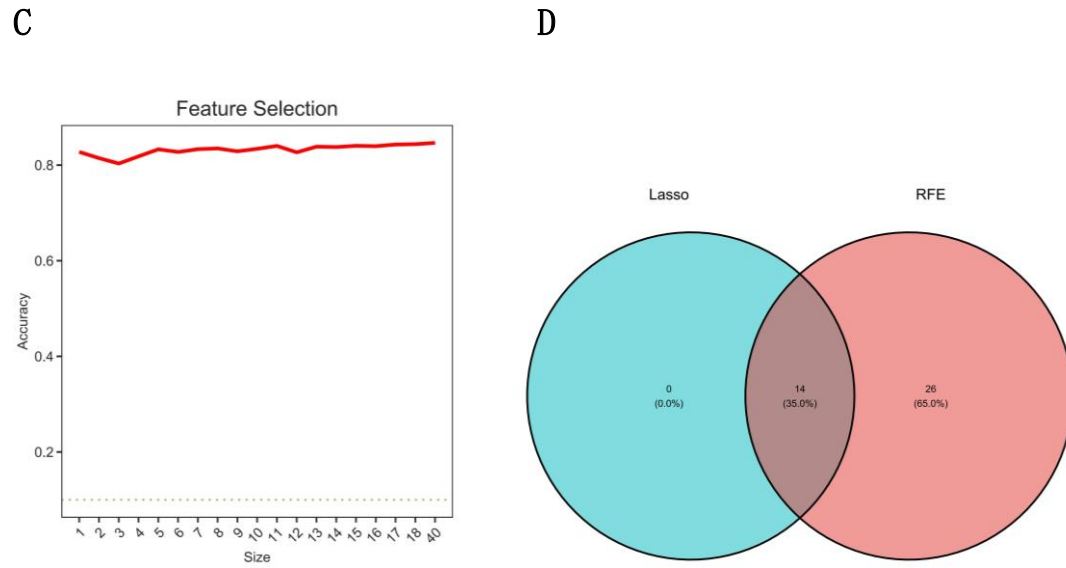


Figure 2 A. Lasso coefficient path plot for 40 variables. B. Cross-validation curve (k-fold cross-validation). C. Feature selection based on recursive elimination. D. Venn diagram of feature selection by recursive elimination and Lasso regression.

3.3 model performance

We constructed 8 machine learning models to predict the 30-day mortality risk in patients with intracerebral hemorrhage. They demonstrated excellent discrimination on the ROC curve, with LR (Logistic Regression) performing the best. LR achieved an AUC greater than 0.8 in both the training and validation sets (Training set: 0.814, 95% CI: 0.802-0.825; Internal validation set: 0.811, 95% CI: 0.785-0.837; External validation set: 0.803, 95% CI: 0.781-0.825) (Figure 3A, B, C).

Table 2 and Table 3 present the detailed performance metrics of the validation sets for the eight models. While the models exhibit varying performance across different metrics, model stands out, achieving high F1 scores and accuracy in both validation sets.

Figure 4A, B, and C show the calibration curves of the eight models on the training and validation sets, providing a critical basis for prediction reliability; the calibration curves of the LR model align relatively well with the ideal line across all three datasets.

Figure 5A, B, and C display the calibration curves of the models across different datasets, and all models demonstrate high net benefits.

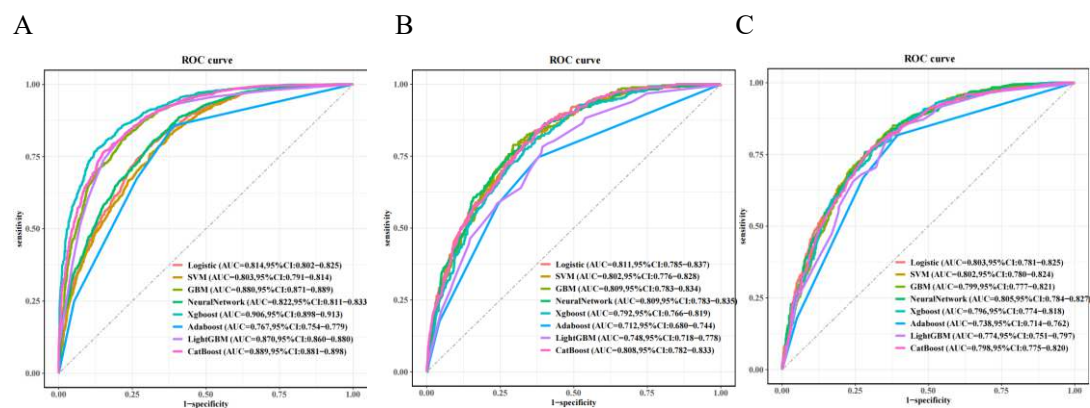


Figure 3A. ROC curve of the training set. **B.** ROC curve of the internal validation set. **C.** ROC curve of the external validation set.

Table 2 Performance of the machine model in predicting 30-day mortality risk in patients with intracerebral hemorrhage in the internal validation cohort

Model	Accuracy	Sensitivity	Specificity	Precision	F1
Logistic	0.678	0.829	0.647	0.328	0.47
SVM	0.703	0.796	0.684	0.344	0.48
GBM	0.72	0.789	0.705	0.358	0.492
NeuralNetwork	0.724	0.757	0.717	0.358	0.486
Xgboost	0.686	0.775	0.667	0.326	0.459
Adaboost	0.642	0.746	0.62	0.29	0.418
LightGBM	0.637	0.782	0.607	0.292	0.426
CatBoost	0.649	0.857	0.606	0.311	0.457

Table 3 Performance of the machine model in predicting 30-day mortality risk in the external validation cohort of intracerebral hemorrhage patients

Model	Accuracy	Sensitivity	Specificity	Precision	F1
Logistic	0.729	0.75	0.717	0.608	0.671
SVM	0.719	0.802	0.671	0.588	0.679
GBM	0.706	0.851	0.622	0.568	0.682
NeuralNetwork	0.732	0.758	0.717	0.611	0.677
Xgboost	0.718	0.776	0.684	0.59	0.67
Adaboost	0.684	0.818	0.605	0.548	0.656
LightGBM	0.701	0.795	0.646	0.568	0.663
CatBoost	0.721	0.779	0.687	0.593	0.674

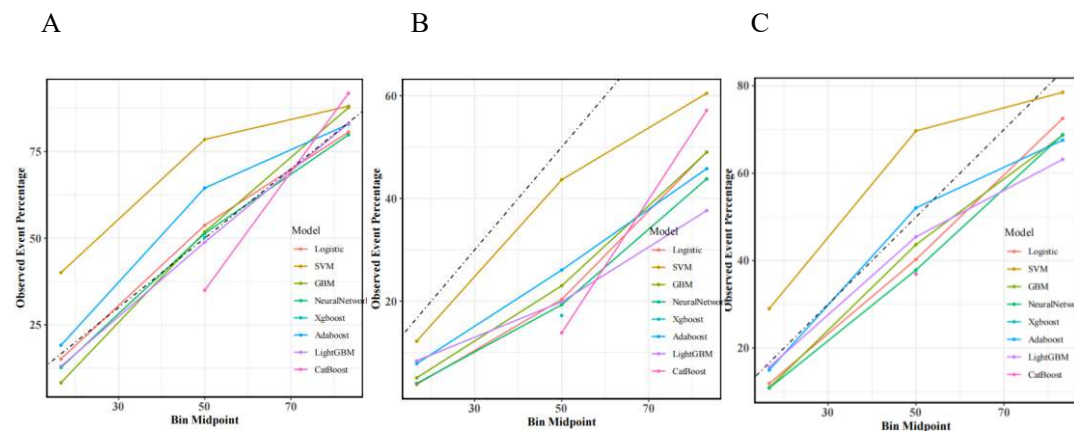


Figure 4A. Calibration curve of the training set. B. Calibration curve of the internal validation set. C. Calibration curve of the external validation set.

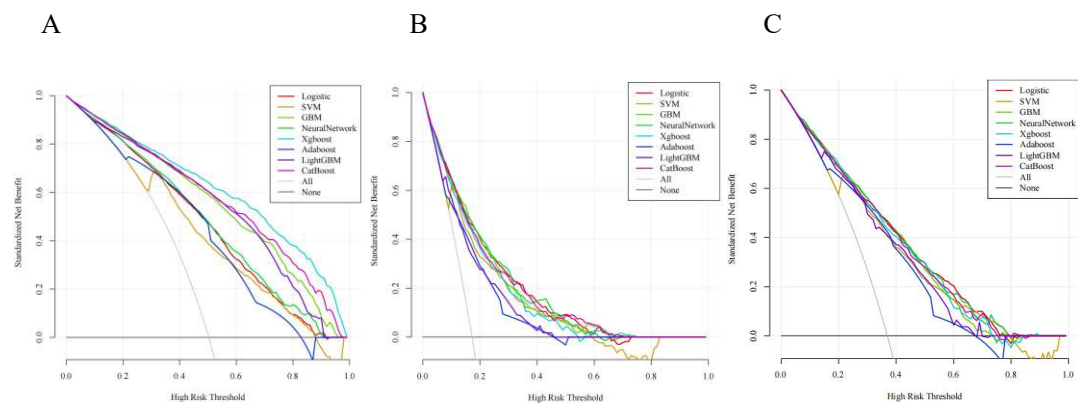


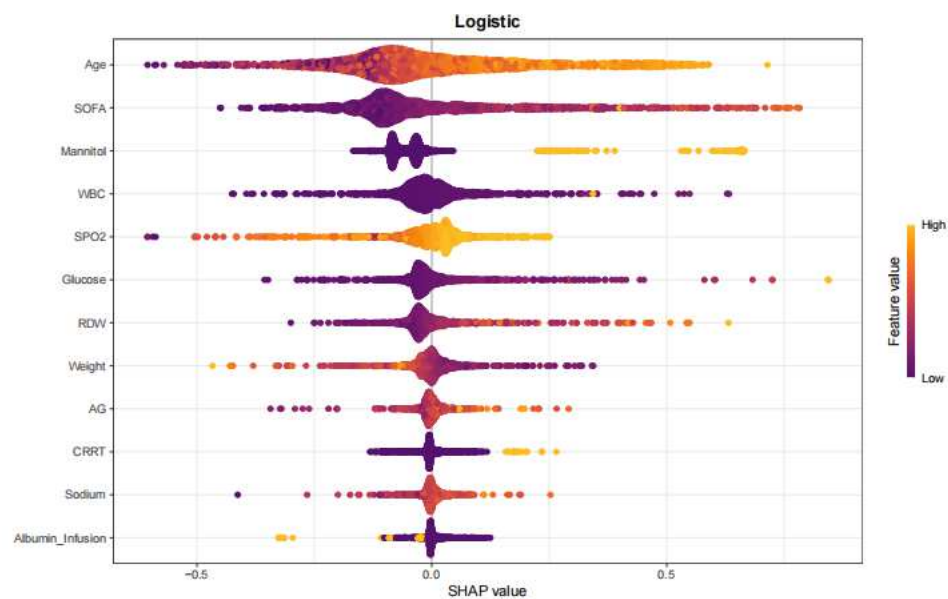
Figure 5A. Clinical decision curve of the training set. B. Clinical decision curve of the internal validation set. C. Clinical decision curve of the external validation set

3.4 interpretability analysis

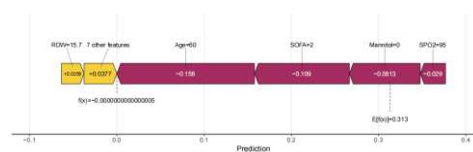
To elucidate the clinical significance of the LR model, we employed the SHAP method to interpret its prediction process and results. Through SHAP analysis, we identified and explained the importance of predictive features to enhance its applicability in clinical decision-making. Two levels of explanation are provided: global explanation describes the overall function of the model, while local explanations pertain to individual prediction results. The SHAP summary plot in Figure 6A illustrates the global explanation. The SHAP beeswarm plot is a visualization tool that intuitively presents the importance of variables and their impact magnitude in a dataset. For the mortality risk prediction model in patients with intracerebral hemorrhage, the feature importance ranking is as follows: age, Sequential Organ Failure Assessment score, mannitol, white blood cell count, oxygen saturation, glucose, red cell distribution width, body weight, anion gap, continuous renal replacement therapy, sodium, and albumin infusion.

Figure 6D, E, and F show the SHAP dependence plots for the Sequential Organ Failure Assessment score, albumin infusion, and age, respectively, intuitively presenting the impact of these three predictors on the LR model's prediction results. The x-axis represents the feature values, and the y-axis shows the SHAP values, which quantify the significant impact of each feature on the prediction results: increased age and SOFA score are associated with an increased risk of death. Albumin infusion can reduce the risk of death. This chart clearly displays the positive or negative effects of the predictors on the prediction results: positive SHAP values indicate an increased risk of death, while negative values mean a reduced risk. Fig 6G shows the top 12 feature variables ranked by their importance to the prediction results.

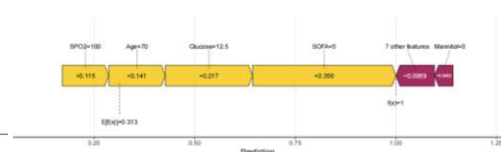
A



B



C



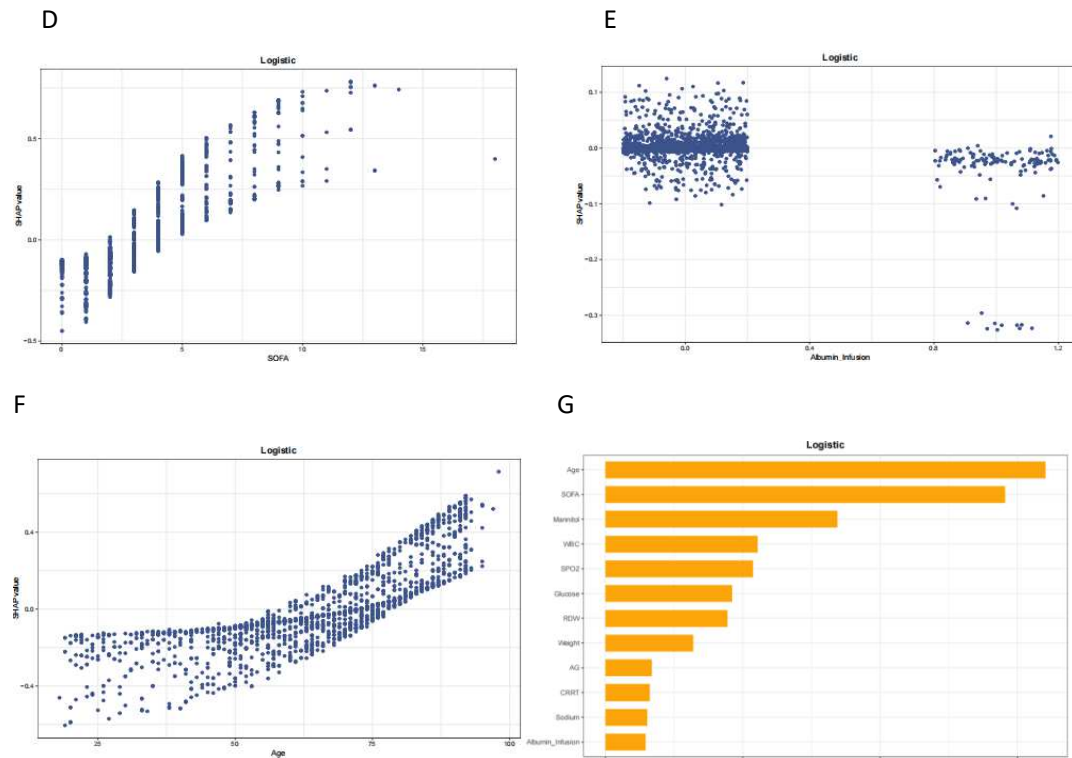


Figure 6 Interpretability analysis of the LR model. A. SHAP importance plot. B. C. D. SHAP dependence plots for Sequential Organ Failure Assessment score. E. SHAP dependence plot for albumin infusion. F. SHAP dependence plot for age. G. Feature importance ranking plot of the LR regression model features

4. Discussion

In this study, we constructed eight machine learning models to predict the 30-day mortality risk of intracerebral hemorrhage. By comparing the performance of each model on the training set, internal validation set, and independent external test set (AUC, accuracy, calibration curves, etc.), we found that the Logistic Regression (LR) model was superior. Furthermore, this study innovatively incorporated post-albumin infusion as a new feature variable and found it can reduce the mortality of patients with intracerebral hemorrhage. Additionally, SHAP was used to interpret the LR model. The SHAP analysis showed that the importance of features contributing to the

model was ranked as follows: age, Sequential Organ Failure Assessment (SOFA) score, mannitol, white blood cell count, oxygen saturation, glucose, red cell distribution width (RDW), body weight, anion gap, continuous renal replacement therapy, sodium, and albumin infusion. This is consistent with the findings of the Dunham study, which showed a significant increase in mortality in patients aged ≥ 70 years, and the results of multiple other researchers also indicate that increasing age is associated with in-hospital mortality^[7,8,9,10]. The SOFA score, as the gold standard for predicting sepsis mortality^[11], indicates more severe dysfunction in various systems with higher scores, often predicts poor patient outcomes. Mannitol, as a hyperosmotic solution, can be used to reduce fluid volume in the brain. Mannitol causes osmotic water movement from the brain parenchyma, resulting in reduced brain water content and increased extracellular volume. Reducing brain water content can alleviate perilesional edema. Consistent with previous studies, the use of mannitol has a positive effect on the clinical prognosis of patients with severe traumatic brain injury^[12]. White blood cell (WBC) count refers to the number of blood cells per unit volume, and its elevation is usually associated with acute infection, uremia, severe burns, and other conditions^[13]. WBC count has long been used in clinical settings to reflect an individual's immune status. Previous studies have also confirmed that sepsis is closely related to the systemic inflammatory response, so there is a certain theoretical basis for reflecting this indicator to evaluate the prognosis of infected patients^[14]. This study found that patients with higher oxygen saturation had higher mortality. Previous studies have found an association between mortality and median blood oxygen levels,

with both hypoxemia and hyperoxemia being associated with higher mortality risks. The relationship between pulse oximetry oxygen saturation (SpO₂) and in-hospital mortality in hospitalized patients was U-shaped^[15]. Red cell distribution width (RDW) represents the heterogeneity of red blood cell size, with higher values indicating greater variability. Mildly elevated RDW is associated with oxidative stress, chronic inflammatory responses, and impaired erythropoiesis^[16,17]. Inflammatory cytokines can affect the survival of circulating red blood cells, reduce their deformability, inhibit their maturation, and lead to increased red blood cell-leukocyte interactions^[18,19]. Excessive oxidative stress reduces blood cell activity^[20]. Furthermore, elevated RDW levels are associated with the risk of recurrent small artery occlusion^[21]. Pathological changes in small cerebral vessels can trigger small artery occlusion and involve deep intracerebral hemorrhage^[22, 23].

Hyperglycemia exacerbates brain injury by increasing intracellular osmotic pressure, leading to cerebral edema and neuronal dysfunction. Due to metabolism, the production of large amounts of amino acids, nitric oxide (NO), and nitrites contributes to brain swelling. Insulin resistance induced by spontaneous intracerebral hemorrhage (ICH), combined with sustained hyperglycemia, promotes endothelial damage and vascular rupture, further aggravating hematoma expansion. Long-term ischemic events can induce anaerobic metabolism and metabolite accumulation, exacerbating cerebral edema^[24,25,26,27]. This study found that underweight patients have a significantly increased risk of mortality. Obesity may serve as a reserve against post-stroke metabolic and energy imbalance. Stroke triggers various acute

pathophysiological responses, most of which lead to enhanced catabolism uninhibited by anabolism ^[28]. With metabolic imbalance, coupled with prolonged bed rest, lack of activity, and paralysis, weight loss accelerates, leading to poor post-stroke outcomes, while fat mass may exert a protective role to prevent this vicious cycle ^[29].

Serum anion gap (AG) represents the difference between unmeasured positive and negative ions in the blood, serving as a critical indicator for determining the type of metabolic acidosis and assessing acid-base status ^[30,31]. Numerous studies have demonstrated a linear association between elevated anion gap (AG) levels and increased mortality in patients admitted to the intensive care unit (ICU) with conditions such as cerebrovascular accidents ^[32], aortic aneurysms ^[33], acute kidney injury (AKI) ^[34], acute pancreatitis, and sepsis ^[35]. Hypernatremia is the mortality risk in critically ill patients ^[36,37,38]. It is a key factor in morbidity and mortality among patients with neurological diseases. In neurosurgical intensive care units, the incidence of hypernatremia in patients with severe head injury is 36.9%, and it is independently associated with early mortality. Hyperemia is more prevalent in patients with severe head injury who also have central diabetes insipidus and lower Glasgow Coma Scale scores ^[39]. Albumin exerts neuroprotective effects on the occurrence and development of ischemic stroke, with mechanisms encompassing antioxidant, anti-inflammatory, and anti-edema functions, maintaining vascular permeability, and regulating glial cell metabolism ^[40,41]. It is widely applied in the clinical treatment of intracerebral hemorrhage. This study efficiently screened variables using Lasso regression and recursive elimination, simultaneously reducing

the number of variables, which effectively improved the clinical applicability of the parameters in the model are common clinical indicators that are easy to obtain, facilitating the clinical application and promotion of the model. The machine learning model can serve as a screening tool for high-risk patients, providing auxiliary information for clinical decision-making. This study utilized the SHAP method to conduct global and local interpretations of the machine learning model, enhancing its visualization and transparency. The model developed in this study demonstrates significant discriminative performance and application value in predicting.

This study is not without limitations. First, this is a retrospective study, and selection bias is inevitable. Further prospective studies are needed to elevate the level evidence to support our findings. Second, although the SHAP method enhances model interpretability, interpreting machine learning models remains challenging. The decision-making process of the model may still be unclear. Therefore, our next step will involve integrating the predictive model into digital health record systems to predict individual cases by extracting patient information and metrics.

5.Conclusion

This study found that albumin infusion reduces the risk of death in patients with intracerebral hemorrhage and successfully developed a prediction tool based on a logistic regression model to predict the 30-day mortality risk in patients with intracerebral hemorrhage combined with albumin infusion characteristics.

6.Declarations

Ethics,funding and Consent to Participate declarations: not applicable

7.References

- [1] Wan Y, Holste K G, Hua Y, et al. Brain edema formation and therapy after intracerebral hemorrhage[J]. Neurobiology of disease, 2023, 176: 105948.
- [2] Tsao C W, Aday A W, Almarzooq Z I, et al. Heart disease and stroke statistics—2022 update: a report from the American Heart Association[J]. Circulation, 2022, 145(8): e153-e639.
- [3] Emerson T J. Unique features of albumin: a brief review[J]. Critical care medicine, 1989, 17(7): 690-694.
- [4] Belayev L, Liu Y, Zhao W, et al. Human albumin therapy of acute ischemic stroke: marked neuroprotective efficacy at moderate doses and with a broad therapeutic window[J]. Stroke, 2001, 32(2): 553-560.
- [5] Li X Y, Chen W S, Qu Z K, et al. Early use of albumin may increase the risk of sepsis-associated acute kidney injury in sepsis patients: a target trial emulation[J]. Military Medical Research, 2025, 12(1): 51-75
- [6] Bako A T, Vahidy F S. Machine Learning-Based Models for Intracerebral Hemorrhage In-Hospital Mortality Prediction[J]. Journal of the American Heart Association, 2025, 14(5): e039398.
- [7] Bruijns S R, Guly H R, Bouamra O, et al. The value of traditional vital signs,

- shock index, and age-based markers in predicting trauma mortality[J]. *Journal of Trauma and Acute Care Surgery*, 2013, 74(6): 1432-1437.
- [8] Lee J H, Kim M J, Hong J Y, et al. The elderly age criterion for increased in-hospital mortality in trauma patients: a retrospective cohort study[J]. *Scandinavian journal of trauma, resuscitation and emergency medicine*, 2021, 29(1): 133-16.
- [9] Campbell-Furtick M, Moore B J, Overton T L, et al. Post-trauma mortality increase at age 60: a cutoff for defining elderly?[J]. *The American Journal of Surgery*, 2016, 212(4): 781-785.
- [10] Cooper D, Jacobson M, Bart R S. Hypopigmented macules[J]. *Archives of dermatology*, 1992, 128(9): 1269-1270.
- [11] Lu J, Dong Z, Ye L, et al. Predictive value of SOFA, PCT, Lactate, qSOFA and their combinations for mortality in patients with sepsis: A systematic review and meta-analysis[J]. *Plos one*, 2025, 20(9): e0332525.
- [12] Syahrul S, Musadir N, Hidayaturrahmi H, et al. Effectiveness of Mannitol use on clinical outcomes of severe traumatic brain injury patients[J]. *F1000Research*, 2024, 13: 548.
- [13] Legrand M, Tolwani A. Anticoagulation strategies in continuous renal replacement therapy[C]//*Seminars in dialysis*. 2021, 34(6): 416-422.
- [14] Li Y, Deng X, Feng J, et al. Predictors for short-term successful weaning from continuous renal replacement therapy: a systematic review and meta-analysis[J]. *Renal Failure*, 2023, 45(1): 2176170.

- [15]Yin H, Yang R, Xin Y, et al. In-hospital mortality and SpO₂ in critical care patients with cerebral injury: data from the MIMIC-IV Database[J]. BMC anesthesiology, 2022, 22(1): 386-396.
- [16]Salvagno G L, Sanchis-Gomar F, Picanza A, et al. Red blood cell distribution width: a simple parameter with multiple clinical applications[J]. Critical reviews in clinical laboratory sciences, 2015, 52(2): 86-105.
- [17]Wang Q, Zennadi R. Oxidative stress and thrombosis during aging: the roles of oxidative stress in RBCs in venous thrombosis[J]. International journal of molecular sciences, 2020, 21(12): 4259-5267.
- [18]Morcneau F, Dicato M, Diederich M. Pro-inflammatory cytokine-mediated anemia: Regarding molecular mechanisms of erythropoiesis[J]. Mediators of inflammation, 2009, 2009(1): 405-516.
- [19]Li N, Zhou H, Tang Q. Red blood cell distribution width: a novel predictive indicator for cardiovascular and cerebrovascular diseases[J]. Disease markers, 2017, 2017(1): 7089493.
- [20]Wang Q, Zennadi R. Oxidative stress and thrombosis during aging: the roles of oxidative stress in RBCs in venous thrombosis[J]. International journal of molecular sciences, 2020, 21(12): 4259.
- [21]Guan Y, Zuo W, Jia K, et al. Association of red blood cell distribution width with stroke prognosis among patients with small artery occlusion: a hospital-based prospective follow-up study[J]. International Journal of General Medicine, 2022, 15: 7449.

- [22] Van Der Meijden P E J, Heemskerk J W M. Platelet biology and functions: new concepts and clinical perspectives[J]. *Nature Reviews Cardiology*, 2019, 16(3): 166-179.
- [23] Liang H, Liu P, Guo L, et al. Predictive value of admission red cell distribution width-to-platelet ratio for 30-day death in patients with spontaneous intracerebral hemorrhage: An analysis of the MIMIC database[J]. *Frontiers in Neurology*, 2023, 14: 1221335.
- [24] Aleman L, Guerrero J. Hiperglicemia por sepsis: del mecanismo a la clínica[J]. *Revista médica de Chile*, 2018, 146(4): 502-510.
- [25] Scheen M, Giraud R, Bendjelid K. Stress hyperglycemia, cardiac glucotoxicity, and critically ill patient outcomes current clinical and pathophysiological evidence[J]. *Physiological reports*, 2021, 9(2): e14713.
- [26] Marik P E, Bellomo R. Stress hyperglycemia: an essential survival response![J]. *Critical care*, 2013, 17(2): 305-315.
- [27] Gong Y, Wang Y, Chen D, et al. Predictive value of hyperglycemia on prognosis in spontaneous intracerebral hemorrhage patients[J]. *Heliyon*, 2023, 9(3)234-2510
- [28] Scherbakov N, Dirnagl U, Doehner W. Body weight after stroke: lessons from the obesity paradox[J]. *Stroke*, 2011, 42(12): 3646-3650.
- [29] Wakisaka K, Matsuo R, Matsumoto K, et al. Non-linear association between body weight and functional outcome after acute ischemic stroke[J]. *Scientific reports*, 2023, 13(1): 8697-8723.

- [30] Fenves A Z, Emmett M. Approach to patients with high anion gap metabolic acidosis: core curriculum 2021[J]. American Journal of Kidney Diseases, 2021, 78(4): 590-600.
- [31] Kraut J A, Nagami G T. The serum anion gap in the evaluation of acid-base disorders: what are its limitations and can its effectiveness be improved?[J]. Clinical Journal of the American Society of Nephrology, 2013, 8(11): 2018-2024.
- [32] Liu X, Feng Y, Zhu X, et al. Serum anion gap at admission predicts all-cause mortality in critically ill patients with cerebral infarction: evidence from the MIMIC-III database[J]. Biomarkers, 2020, 25(8): 725-732.
- [33] Gao Y, Hong Z, Shen R, et al. Association between anion gap and mortality of aortic aneurysm in intensive care unit after open surgery[J]. BMC Cardiovascular Disorders, 2021, 21(1): 458-469.
- [34] Zhong L, Xie B, Ji X W, et al. The association between albumin corrected anion gap and ICU mortality in acute kidney injury patients requiring continuous renal replacement therapy[J]. Internal and Emergency Medicine, 2022, 17(8): 2315-2322.
- [35] Zhu Y, He Z, Jin Y, et al. Serum anion gap level predicts all-cause mortality in septic patients: a retrospective study based on the MIMIC III database[J]. Journal of Intensive Care Medicine, 2023, 38(4): 349-357.
- [36] Annane D, Bellissant E, Bollaert P E, et al. Corticosteroids for treating sepsis[J]. Cochrane database of systematic reviews, 2015 (12):12-23.
- [37] Annane D, Briegel J, Granton D, et al. Corticosteroids for treating sepsis in

children and adults[J]. Cochrane Database of Systematic Reviews, 2025 (6)34-41.

[38]Fang F, Zhang Y, Tang J, et al. Association of corticosteroid treatment with outcomes in adult patients with sepsis: a systematic review and meta-analysis[J]. JAMA internal medicine, 2019, 179(2): 213-223.

[39]Qi Z, Lu J, Liu P, et al. Nomogram prediction model of hypernatremia on mortality in critically ill patients[J]. Infection and Drug Resistance, 2023: 143-153.

[40]Belayev L, Liu Y, Zhao W, et al. Human albumin therapy of acute ischemic stroke: marked neuroprotective efficacy at moderate doses and with a broad therapeutic window[J]. Stroke, 2001, 32(2): 553-560.

[41]Li X Y, Chen W S, Qu Z K, et al. Early use of albumin may increase the risk of sepsis-associated acute kidney injury in sepsis patients: a target trial emulation[J]. Military Medical Research, 2025, 12(1): 51-62.