ORIGINAL RESEARCH

Interpretable Machine Learning Model for Predicting the Prognosis of Guillain-Barré Syndrome Patients

Junshuang Guo[1,2], Ruike Zhang[1], Ruirui Dong[1], Fan Yang[1], Yating Wang[1], Wang Miao 🆔[1]

[1]Neuro-Intensive Care Unit of the First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan Province, People's Republic of China; [2]Department of Immunology, School of Basic Medical Science, Central South University, Changsha City, Hunan Province, People's Republic of China

Correspondence: Wang Miao, Neuro-Intensive Care Unit of the First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan Province, 450000, People's Republic of China, Email miaowang7211@126.com

Background: Machine learning (ML) is increasingly used in medical predictive modeling, but there are no studies applying ML to predict prognosis in Guillain-Barré syndrome (GBS).

Materials and Methods: The medical records of 223 patients with GBS were analyzed to construct predictive models that affect patient prognosis. Least Absolute Shrinkage and Selection Operator (LASSO) was used to filter the variables. Decision Trees (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost), k-nearest Neighbour (KNN), Naive Bayes (NB), Neural Network (NN). Light Gradient Boosting Machine (LGBM) and Logistic Regression (LR) were used to construct predictive models. Clinical data from 55 GBS patients were used to validate the model. SHapley additive explanation (SHAP) analysis was used to explain the model. Single sample gene set enrichment analysis (ssGSEA) was used for immune cell infiltration analysis.

Results: The AUCs (area under the curves) of the 8 ML algorithms including DT, RF, XGBoost, KNN, NB, NN, LGBM and LR were as follows: 0.75, 0.896 0.874, 0.666, 0.742, 0.765, 0.869 and 0.744. The accuracy of XGBoost (0.852) was the highest, followed by LGBM (0.803) and RF (0.758), with F1 index of 0.832, 0.794, and 0.667, respectively. The results of the validation set data analysis showed AUCs of 0.839, 0.919, and 0.733 for RF, XGBoost, and LGBM, respectively. SHAP analysis showed that the SHAP values of blood neutrophil/lymphocyte ratio (NLR), age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus nerve were 0.821, 0.645, 0.517, 0.401 and 0.109, respectively.

Conclusion: The combination of NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus used to predict short-term prognosis in patients with GBS has a good predictive value.

Keywords: Guillain-Barré syndrome, machine learning, prognosis, SHAP

Introduction

Guillain-Barré syndrome (GBS) is an autoimmune-mediated peripheral neuropathy with acute onset, flaccid paralysis of the limbs, and electromyography showing demyelination or axonal damage to peripheral nerves. GBS is often secondary to infections with pathogens such as Campylobacter jejuni, cytomegalovirus, and Epstein-Barr virus. Approximately 100,000 people develop the disease each year, and 20–30% of patients develop respiratory failure during the course of the disease requiring ventilator-assisted ventilation. Treatment is mainly based on intravenous immunoglobulin (IVIG) and plasma exchange (PE).[1] Although the existing clinical treatment regimens have significant therapeutic effects on some patients, some patients still require mechanical ventilation during the acute progression of the disease, and some patients are left with severe sequelae such as decreased muscle strength and sensory abnormalities, which seriously affects the quality of life of the patients and aggravates the economic burdens of the family and the society.[2,3] More seriously, some patients have a rapid progression and severe condition, and the above treatment cannot significantly improve their condition. A double blind, randomized, placebo controlled trial does not provide evidence that patients with GBS with a poor diagnosis benefit from a second IVIG course, moreover, it enters a risk of series adverse events.[4] Therefore, studying the influencing factors of patient prognosis and early intervention is crucial. The study of the factors

affecting the prognosis of patients and early intervention can alleviate the symptoms of patients and reduce medical expenses.

With the advent of the big data era, machine learning (ML) is increasingly used in medical predictive modelling, and some studies have demonstrated that ML methods outperform traditional methods in disease prediction.[5–8] The methods of ML each have their own advantages. The structure of the decision tree (DT) is intuitive, Random Forest (RF) improves the prediction accuracy of the model through integrated learning, Extreme Gradient Boosting (XGBoost) usually yields high accuracy when dealing with complex datasets, k-nearest Neighbour (KNN) performs well on small datasets, Naive Bayes (NB) is very fast in both training and prediction, Neural Network (NN) is suitable for complex tasks, Light Gradient Boosting Machine (LGBM) dramatically improves the speed of training, and Logistic Regression (LR) models are simple and easy to implement and understandable. SHapley additive explanations (SHAP) is an approach to address the interpretability of models. By calculating the SHAP value, it reflects the marginal contribution of features to the model output. SHAP analysis is suitable for explaining the model performance,[9] which can visualise the prediction direction of each variable to the model results as well as the contribution of each variable, and also show the contribution of each variable to the prediction of a single patient through the force plot.

Previous studies have used logistic regression to study the factors affecting the prognosis of patients with GBS.[10–12] However, there are no studies applying ML to predict the early prognosis of GBS. Early intervention for modifiable risk factors can help to improve the prognosis of patients to reduce the economic burden. Considering the superiority of ML algorithms, eight methods of ML were used to construct a prediction model to study the influencing factors affecting the short-term prognosis of GBS patients at the time of discharge from the hospital, so as to provide a reference for clinical treatment and to improve the prognosis of patients.

Materials and Methods
Data Collection
In this study, the clinical data containing 278 patients with GBS were collected from the Department of Neurology of the First Affiliated Hospital of Zhengzhou University. Two hundred and twenty-three patients admitted between October 2019 and December 2021 were classified as the training set and 55 patients admitted between January 2023 and June 2023 were classified as the validation set. Given that a bedridden state and the need for assisted ventilation more accurately reflect the severity of the disease, this study aimed to examine patients with more severe disease, as indicated by the Hughes functional grading scale (HFGS) of $\geq$ 4 points. The patients in the training and validation sets were divided into a good prognosis group (HFGS<4) and a poor prognosis group (HFGS$\geq$4) according to whether the patients had a HFGS score was less than 4 at the time of discharge, and the factors affecting the prognosis of GBS patients were analysed.

Variable Screening
Univariate analyses were performed and variables with P value < 0.05 were included in the Least Absolute Shrinkage and Selection Operator (LASSO) for variable screening to avoid overfitting, and λ was adjusted by the tenfold cross-validation method.

Model Construction, Evaluation, Validation and Model Interpretability Analysis
DT, RF, XGBoost, KNN, NB, NN, LGBM and LR which are the eight ML methods to build the prediction model with sampling set to bootstrap=1000. The accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1 index were calculated to compare the above eight ML models. This study used the bootstrap method (with 1000 samples placed back) in model construction. Multiple studies have proven that bootstrap is an effective model validation method for small sample studies.[13–16] In addition, we collected data that differed from the time period of patient admission in the training set to assess the generalisation ability of the model.

The importance of variables was analyzed to determine the impact of variables on the predicted outcome. The importance of each variable was calculated according to the internal algorithm of the model, and the importance of

variables is visualized. The receiver operation characteristic curve (ROC) was drawn, and the area under curve (AUC) was used to evaluate the discriminant performance of different models, and 95% CI was calculated. Decision curve analysis was used to evaluate the clinical applicability of the model. SHAP analysis were used to explain the model, calculate SHAP values and sort them. The force plot was used to explain the single sample, showing the direction and degree of the influence of the variable value of a single individual on the outcome. The interactive SHAP graph was drawn to show the influence of the interaction between the two variables.

Immune Infiltration Analysis

Considering that GBS is an immune-related disease, immune factors are involved. In order to explore the immune cell infiltration of GBS patient samples, this study used single sample gene set enrichment analysis (ssGSEA) for immune cell infiltration analysis. This method estimates the relative enrichment of the gene set in the sample by comparing the gene expression data of each sample with a specific gene set (immune cell gene set). This study analyzed the GSE31014 dataset (Figure 1).

Statistical Analysis

SPSS 25.0 and R software (version 4.0.2, Vienna, Austria) were used for data analysis. Continuous variables were compared using t test or Mann–Whitney u-test, and categorical variables were analyzed using chi-square test or Fisher's exact test. P value < 0.05 was considered statistically significant.

Results

Results of Univariate Analyses

The demographic and clinical characteristics of the training and validation sets are shown in Table S1. In the training set, there were 121 patients in the good prognosis group, of which the number of patients with HFGS scores of 1, 2, and 3 at discharge were 1, 34, and 86, respectively. There were 102 patients in the poor prognosis group, of whom the number of patients with HFGS scores of 4, 5, and 6 at the time of discharge was 89, 12, and 1, respectively. In the training set, compared with patients with good prognosis, patients with poor prognosis were older, the proportion of mechanical ventilation was higher, the proportion of hyporeflexia, the proportion of abnormal glossopharyngeal vagus nerve was higher, white blood cells, neutrophils, neutrophil/lymphocyte ratio (NLR), C-reactive protein, blood glucose, cerebrospinal fluid lymphocytes, cerebrospinal fluid glucose were higher (P < 0.05) (Table 1).

The results of Machine Learning Algorithm

Ten variables with statistical differences in univariate analysis were included in LASSO to further screen variables (Figure 2A). Finally, five variables were obtained (Figure 2B), which were NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus. The model containing the above five variables was constructed by 8 ML methods. The results showed that XGBoost had the highest accuracy and F1 index (0.852, 0.832), followed by LGBM (0.803, 0.794) (Figure 3A–F, Table 2). In the XGBoost, LGBM and RF model, the order of importance of variables is blood NLR, age, mechanical ventilation, hyporeflexia, and abnormal glossopharyngeal vagus nerve (Figure 3G–I).

Model Visualization and Evaluation

The AUCs of the prediction models constructed by DT, RF, XGBoost, KNN, NB, NN, LGBM and LR were 0.75, 0.896, 0.874, 0.666, 0.742, 0.765, 0.869, 0.744, respectively (Figure 4A). AUC values of the model in the validation set were 0.919 (XGBoost), 0.733 (LGBM), and 0.839 (RF) (Figure 4B). The decision curve showed that the prediction model based on the data of this study had a higher upper limit of the prediction threshold probability of poor prognosis (Figure 4C).
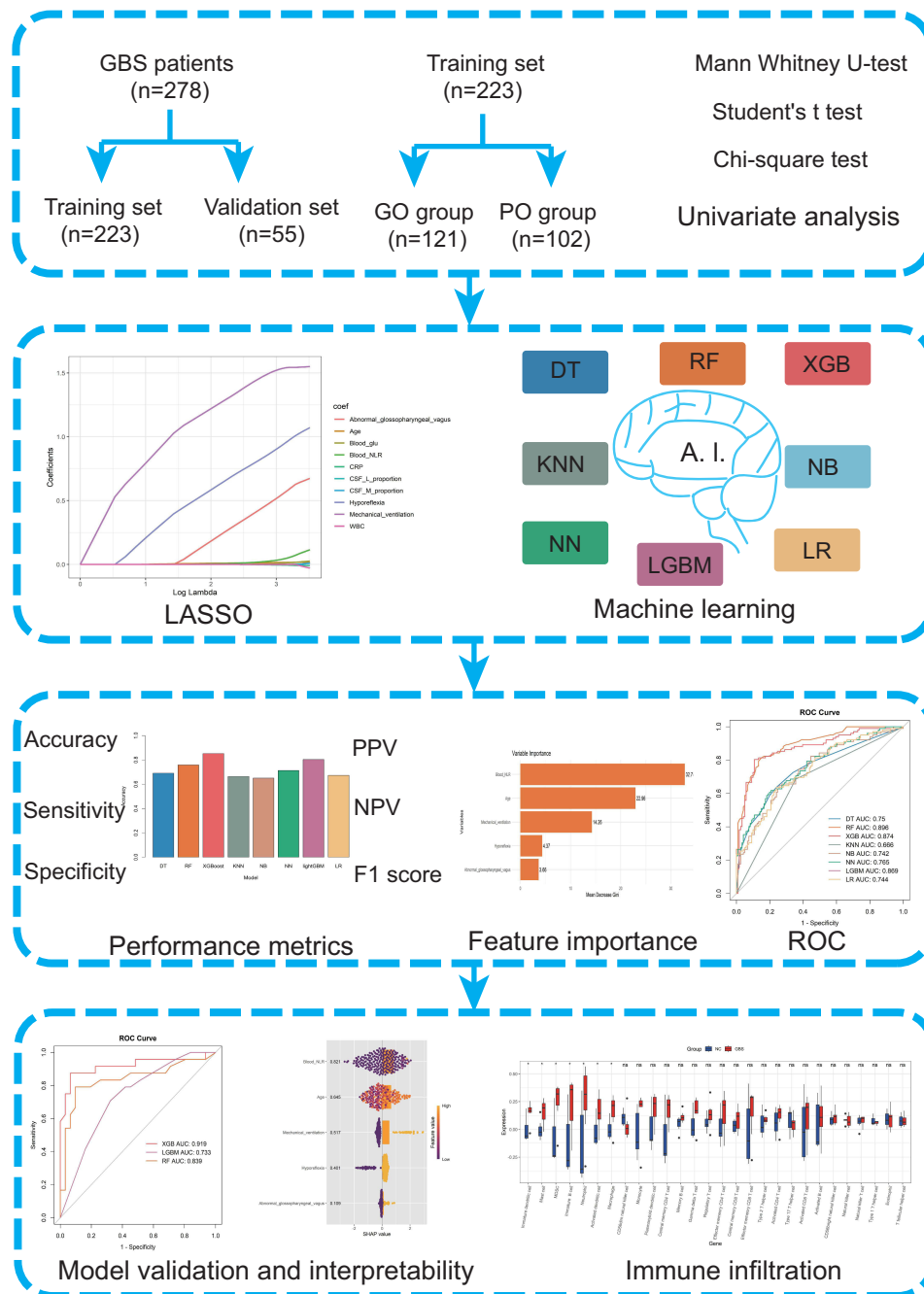
Figure 1 Study design and workflow of patient selection.
Abbreviations: GBS, Guillain - Barré Syndrome, GO, good prognosis, PO, poor prognosis, LASSO, Least Absolute Shrinkage and Selection Operator, DT, Decision Trees, RF, Random Forest, XGBoost, Extreme Gradient Boosting, KNN, k-nearest Neighbour, NB, Naive Bayes, NN, Neural Network, LGBM, Light Gradient Boosting Machine, LR, Logistic Regression, ROC, receiver operation characteristic curve, PPV, positive predictive value, NPV, negative predictive value.

Interpretability of the Model

SHAP analysis showed that the SHAP values of blood NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus nerve were 0.821, 0.645, 0.517, 0.401 and 0.109, respectively (Figure 5A).

The force plot showed that the SHAP benchmark value of the model was 0.19. The output value of the model was 4, five variables all were positive variables, which have the positive impact on poor prognosis. The quantitative effects of

Table 1 Univariate Analysis of Factors Influencing Short-Term Prognosis in Patients with GBS in the Training Set

	Good Prognosis (n=121)	Poor Prognosis (n=102)	x2/t/Z	P
Male (%)	77(63.6)	53(52.0)	3.103	0.078
Age	47.00(26.00–62.00)	55.50(38.75–67.00)	−2.698	0.007
Hospitalization days	14.00(9.50–21.00)	13.00(10.00–25.25)	−0.83	0.406
Diarrhea (%)	26(21.5)	20(19.6)	0.119	0.73
Upper respiratory tract infection (%)	18(14.9)	15(14.7)	0.001	0.972
Mechanical ventilation (%)	7(5.8)	30(29.4)	22.324	<0.001
Hyporeflexia (%)	87(71.9)	91(89.2)	10.301	0.001
Abnormal glossopharyngeal vagus (%)	17(14.0)	26(25.5)	4.654	0.031
Abnormal hypoglossal nerve (%)	0(0.0)	1(1.0)	1.192	0.457
GC (%)	22(18.2)	22(21.6)	0.401	0.527
IVIG (%)	94(77.7)	81(79.4)	0.098	0.755
PE (%)	12(9.9)	19(18.6)	3.508	0.061
Blood test				
WBC (10^9/L)	6.85(5.64–8.64)	7.85(5.76–10.71)	−2.59	0.010
Neutrophil (10^9/L)	4.45(3.23–6.06)	5.57(3.78–8.41)	−3.071	0.002
Lymphocyte (10^9/L)	1.73(1.31–2.12)	1.46(1.01–2.00)	−2.642	0.008
NLR	2.61(1.72–4.16)	3.59(2.22–7.49)	−3.654	<0.001
Monocyte (10^9/L)	0.46(0.38–0.61)	0.53(0.40–0.61)	−1.729	0.084
Blood glucose (mmol/L)	4.95(4.47–6.20)	5.94(5.03–7.36)	−3.854	<0.001
Albumin (g/L)	42.10(38.35–45.30)	41.10(38.48–44.03)	−1.779	0.075
Prealbumin (mg/L)	244.41±82.94	228.71±70.19	1.509	0.133
CRP (mg/L)	1.99(0.78–4.84)	3.63(1.10–11.73)	−2.865	0.004
Thyroid Function test				
FT3 (pmol/L)	4.71(4.15–5.39)	4.48(3.93–4.94)	−1.577	0.115
FT4 (pmol/L)	12.54(11.16–14.60)	12.60(10.79–15.83)	−0.557	0.577
TSH (uIU/mL)	2.13(1.33–3.12)	2.00(1.00–3.03)	−1.837	0.066
CSF test				
Leucocyte (10^6/L)	2.00(2.00–6.00)	2.00(2.00–4.00)	−1.8	0.072
Lymphocyte (%)	62.00(56.00–70.00)	66.00(60.00–73.50)	−2.288	0.022
Monocyte (%)	36.00(27.50–40.00)	30.50(23.00–37.00)	−2.663	0.008
Protein (mg/L)	640.60(449.95–1028.20)	614.50(387.10–1105.55)	−0.643	0.520
Glucose (mmol/L)	3.52(3.14–4.14)	3.89(3.29–4.95)	−2.586	0.010
CSF albumin (mg/dL)	42.54(27.30–191.00)	49.85(21.27–193.48)	−0.419	0.675
Serum albumin (mg/dL)	4419.00(4091.00–5074.00)	4383.00(3892.00–29,450.00)	−0.946	0.344
CSF IgG (mg/dL)	11.88(6.29–43.00)	13.60(4.83–51.93)	−0.126	0.900
Serum IgG (mg/dL)	1345.40(978.20–6980.00)	1699.20(1011.50–6475.00)	−0.718	0.473
Albumin quotient (10^-3)	9.56(5.87–15.87)	8.51(4.91–17.62)	−0.608	0.543
IgG quotient (10^-3)	7.08(3.90–12.51)	6.59(3.34–16.53)	−0.018	0.986
IgG index	0.76(0.64–0.97)	0.77(0.55–1.00)	−0.448	0.654

Abbreviations: GC, Glucocorticoids, IVIG, intravenous immunoglobulin, PE, plasma exchange, WBC, white blood cell, NLR, neutrophil/lymphocyte ratio, CRP, C-reactive protein, FT3, Free triiodothyronine, FT4, Free thyroxin, TSH, thyroid-stimulating hormone, CSF, cerebrospinal fluid, IgG, immunoglobulin G.

mechanical ventilation, age, NLR, abnormal glossopharyngeal vagus nerve, and hyporeflexia were 1.77, 1.11, 0.97, 0.198, and 0.141, respectively (Figure 5B).

The diagonal of the interaction diagram showed the relationship between NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus nerve and poor prognosis. Other subdiagrams outside the diagonal showed the effect of pairwise combinations on poor prognosis. The subgraph of the interaction between NLR and age was the widest, indicating that the combination of variables has the greatest impact on the results (Figure 5C).
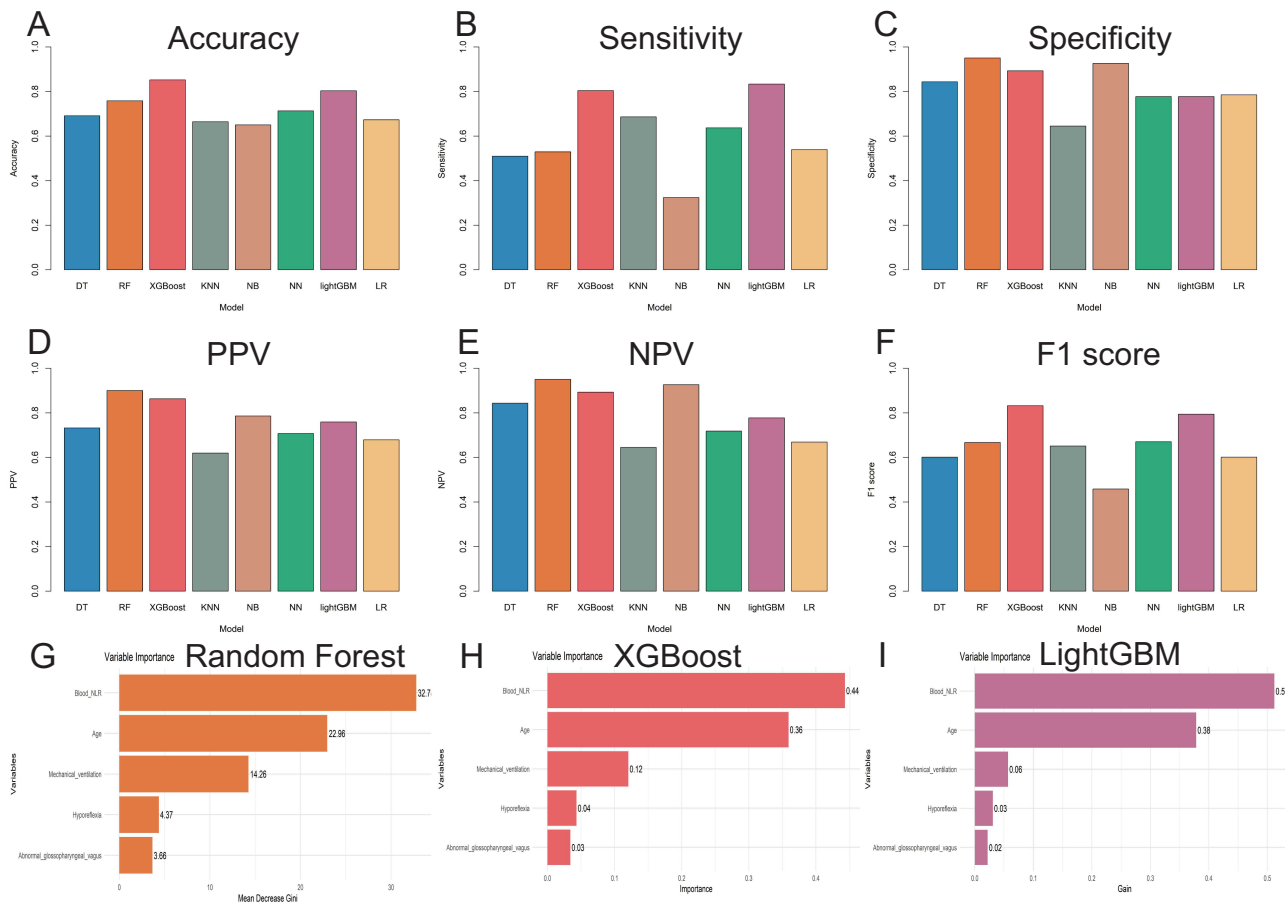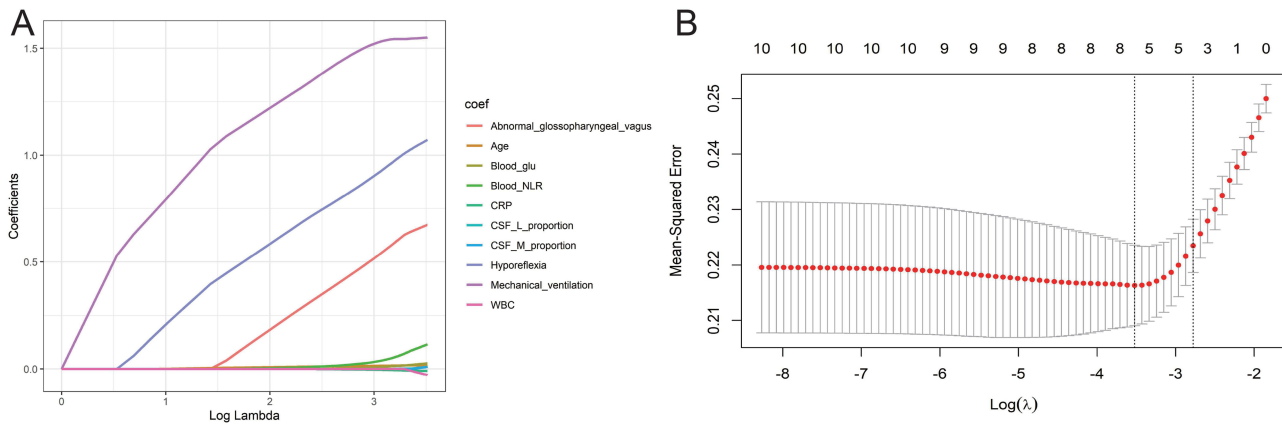
Figure 2 Variable selection and Tenfold cross validation in LASSO algorithm.
Notes: (**A**) Variable selection by LASSO manifests the effect of feature selection, where the coefficients are gradually compressed to zero as lambda increases, (**B**) Tenfold cross-validation of tuning parameter selection in the LASSO.

Figure 3 Model evaluation and variable importance ranking.
Notes: (**A–F**) accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F1 index of the ML models. The model with the highest accuracy (0.852) and F1 index (0.832) are built by the XGBoost algorithm. The specificity (0.950), positive predictive value (0.900), and negative predictive value (0.950) of the model constructed by the RF algorithm are the highest. (**G–I**), RF, XGBoost and LGBM variable importance ranking. In the XGBoost, LGBM and RF model, the order of importance of variables was blood NLR, age, mechanical ventilation, hyporeflexia, and abnormal glossopharyngeal vagus nerve.

Table 2 Performance Analysis of the Models

	AUC	95% CI	Accuracy	PPV	NPV	Sensitivity	Specificity	F1 Index
DT	0.750	0.685–0.815	0.691	0.732	0.843	0.510	0.843	0.601
RF	0.896	0.855–0.936	0.758	0.900	0.950	0.529	0.950	0.667
XGBoost	0.874	0.826–0.923	0.852	0.863	0.893	0.804	0.893	0.832
KNN	0.666	0.603–0.728	0.664	0.619	0.645	0.686	0.645	0.651
NB	0.742	0.677–0.807	0.650	0.786	0.926	0.324	0.926	0.458
NN	0.765	0.703–0.827	0.713	0.707	0.718	0.637	0.777	0.670
LGBM	0.869	0.821–0.916	0.803	0.759	0.777	0.833	0.777	0.794
LR	0.744	0.679–0.808	0.673	0.679	0.669	0.539	0.785	0.601

Abbreviations: AUC, area under the curve, PPV, positive predictive value, NPV, negative predictive value, DT, decision trees, RF, random forest, XGBoost, extreme gradient boosting, KNN, k-nearest neighbor, NB, Naive Bayes, NN, neural network, LGBM, light gradient boosting machine, LR, logistic regression.

Immune Infiltration Analysis

The GSE31014 dataset included 7 GBS patients and 7 normal controls. The results of immune cell infiltration analysis showed that the infiltration levels of immature dendritic cells, mast cells, myeloid-derived suppressor cells, immature
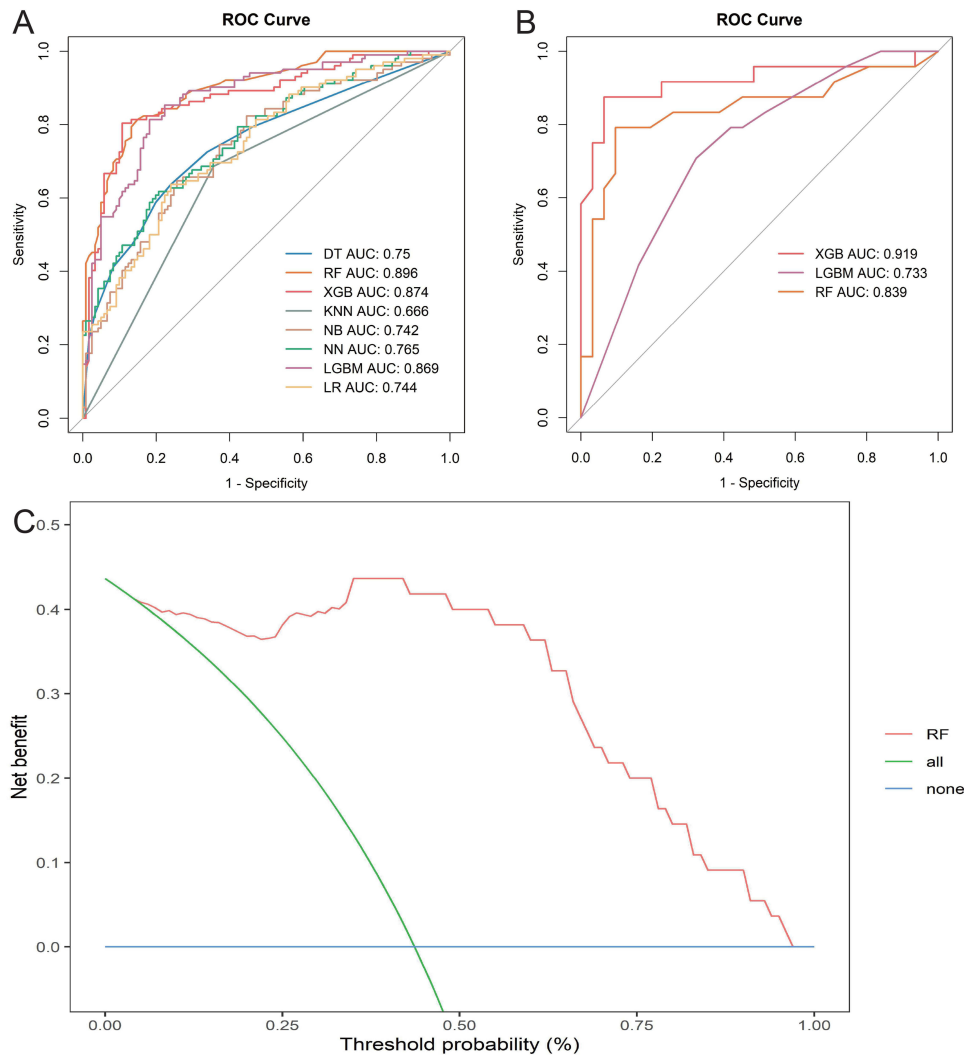
Figure 4 Model visualization and evaluation.
Notes: (**A**) ROC curves in the training set. The model with the highest AUC was constructed using the Random Forest (0.896) algorithm, followed by XGBoost (0.874) and LGBM (0.869). (**B**) AUC of the model in the validation set were 0.919 (XGBoost), 0.733 (LGBM), and 0.839 (RF), (**C**) Decision curve analysis. The net benefit of using the model for clinical decision making was high at all threshold probabilities.
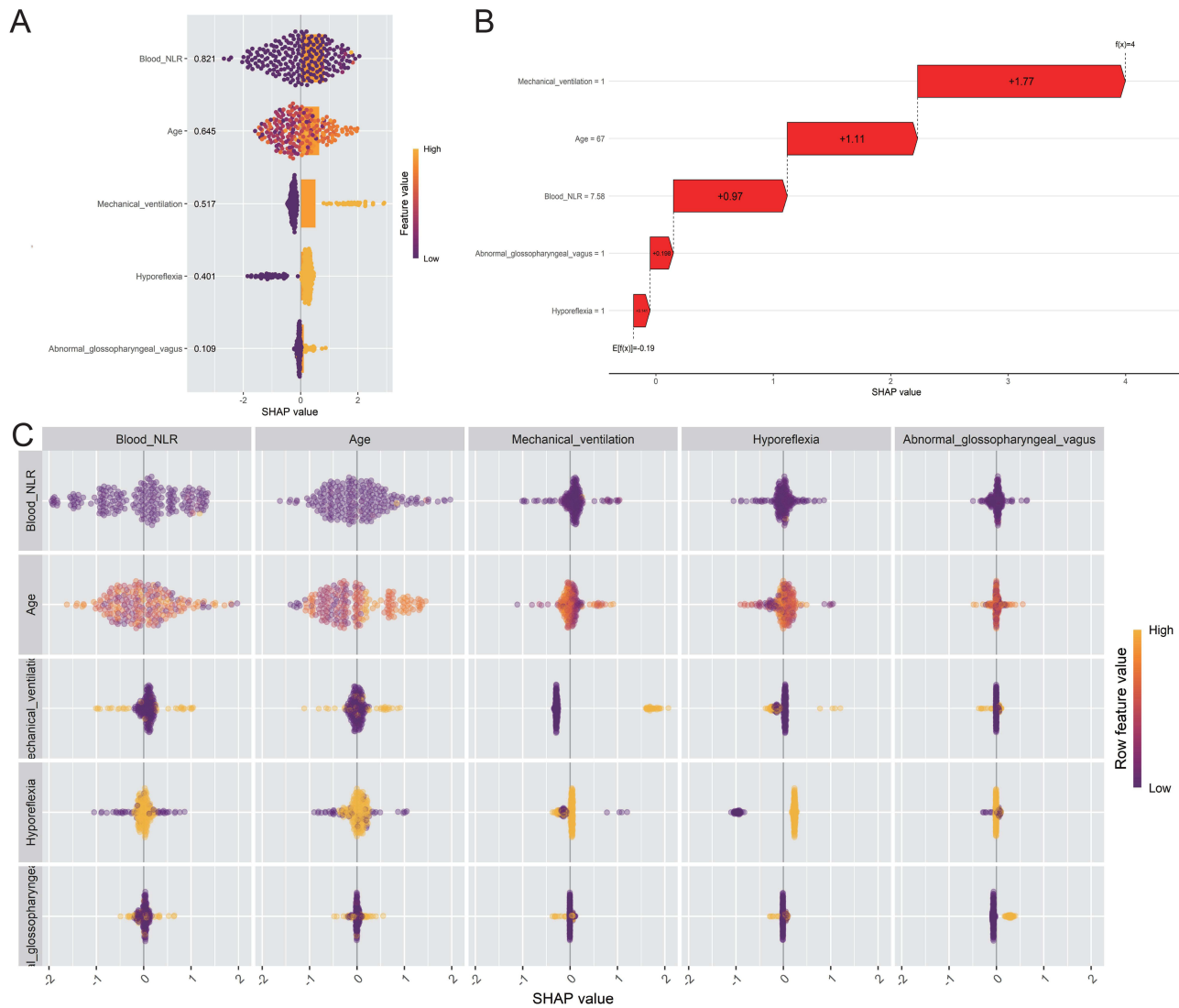
Figure 5 SHAP analysis.
Notes: (**A**) SHAP value analysis. Blood NLR contributed the most to the prediction of the results with a SHAP value of 0.821, followed by age (0.645) and mechanical ventilation (0.517). (**B**) Force plot. The force plot visualized the impact of each feature on each prediction. For this patient, all variables contributed to the poor prognosis. (**C**) Variable pairwise interaction diagram. The SHAP interaction diagram showed the interaction between two variables.

B cells, neutrophils, activated dendritic cells and macrophages in GBS patients were higher than those in the normal control group (P value < 0.05) (Figure 6A). The correlation heatmap of immune infiltration analysis is shown in Figure 6B.

Discussion

Although most patients are treated with PE or IVIG, some patients still have sequelae such as decreased muscle strength and abnormal sensation, and even death. Due to the rapid progression of GBS patients, early identification of the influencing factors affecting the severity of GBS patients and intervention may improve the patient's condition. In this study, a retrospective study of GBS patients was conducted to calculate the differences in clinical symptoms, signs, blood tests, and cerebrospinal fluid tests between the two groups of patients. Eight ML methods were used to construct a prediction model. Due to the integrated learning property of RF, the models have better differentiation and general-ization ability, and the predictive model constructed by the RF algorithm in this study has the best differentiation, with an
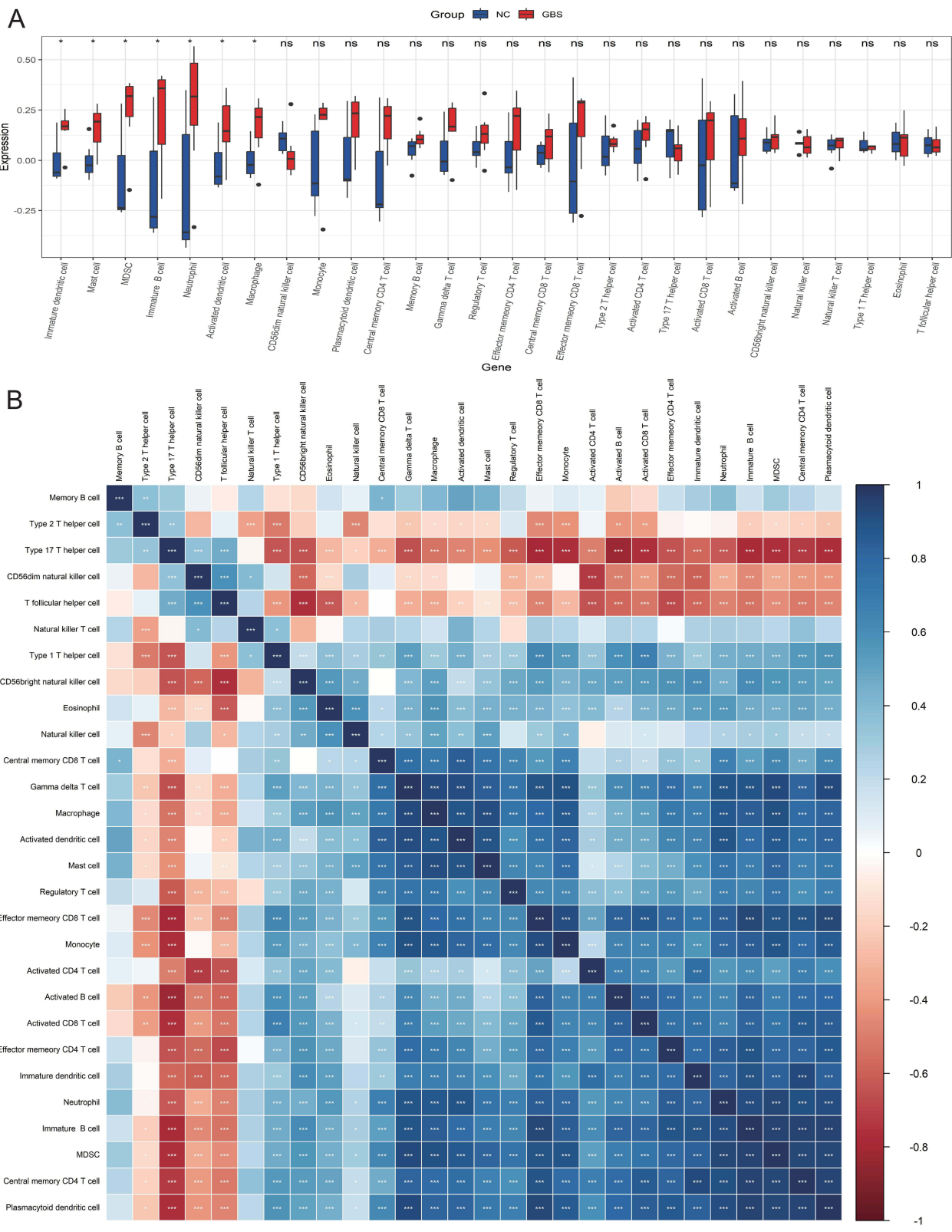
Figure 6 Immunoinfiltration analysis.
Notes: (**A**) Comparison of immune cell infiltration between two groups, higher abundance of neutrophils in GBS patient samples compared to normal controls.
(**B**) Correlation heatmap of immune infiltration analysis, showing the correlation between the abundance of different immune cell types. *: P < 0.05, **: P < 0.01,
***: P < 0.001.
Abbreviation: ns, not significant.

AUC of 0.896. The results showed that NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus nerve can be used to predict the short-term prognosis of GBS patients.

The results show that blood NLR can be used as an indicator to predict the prognosis of AE patients. Univariate analysis showed that the proportion of neutrophils in the poor prognosis group was higher than that in the good prognosis group. The results of the immune infiltration analysis showed that the proportion of neutrophils in GBS patients was higher than that in the control group. GBS is an inflammatory demyelinating disease, and neutrophils increase when inflammatory reactions occur in the body. Neutrophils are prognostic markers for many diseases. Multi-index combination can more accurately reflect the prognosis of patients. In recent years, multi-index joint prediction has been widely used in clinical prediction models, such as NLR and C-reactive protein/albumin ratio (CAR). NLR is a marker of inflammation and infection, and high NLR are independent predictors of poor prognosis in patients with heart disease, cancer, and acute ischaemic stroke.[17–19] The prognostic accuracy of the NLR is higher than traditional indicators of infection such as C-reactive protein or neutrophil and leukocyte counts.[20] In addition, NLR is considered an independent risk factor for the prognosis of ventilator-associated pneumonia.[21] In our study, NLR was an independent risk factor for poor prognosis in patients with GBS.

Previous studies have shown that aging, diarrhea, decreased muscle strength at admission, decreased amplitude of compound muscle action potential[22] and hyponatremia at admission[23] are associated with poor prognosis. The results of this study show that age is a factor affecting the short-term prognosis of patients, which is consistent with previous studies. A number of studies have confirmed that the patient's age is closely related to the severity of the disease.[24,25]

A study of 115 GBS patients showed that patients older than 75 years had a worse prognosis.[26] Meanwhile, a study of 74 very elderly patients over 80 years of age showed that even in the absence of infection, the patient's condition was still severe,[27] and the frequency of requiring a ventilator was significantly higher than that of the control group. A study of 1056 GBS patients found that age can partially determine the progression and severity of GBS patients, and patients younger than 15 years old and older than 56 are more severe.[28] Elderly patients have more basic diseases, poor lung function, and a higher proportion of mechanical ventilation support. In addition, elderly patients have low immunity and are prone to related complications such as infection, which aggravates the disease.

Approximately 30% of patients with GBS develop respiratory failure and therefore require endotracheal intubation and mechanical ventilation support. Respiratory failure is a life-threatening presentation and the leading cause of death in patients with GBS. The results of this study showed that the proportion of mechanical ventilation in the two groups was quite different, and the proportion of mechanical ventilation required during hospitalization in the poor prognosis group was much higher than that in the good prognosis group. A study of 132 patients showed that mechanical ventilation can be used to predict the early prognosis of patients.[29] Besides, a study of 111 Chinese patients showed that mechanical ventilation was a poor prognostic factor.[30] The need for mechanical ventilation indicates that the patient is seriously ill. Although mechanical ventilation can maintain adequate oxygenation and carbon dioxide output, reduce respiratory muscle fatigue and avoid respiratory complications such as atelectasis, however, mechanical ventilation may have some negative effects on patients, such as endotracheal intubation and ventilator-associated pneumonia, especially in the case of prolonged ventilation. Mechanical ventilation may also cause pneumothorax and other injuries. Patients who need mechanical ventilation for a long time are dependent on ventilators and have difficulty in extubation. These complications will aggravate the patient's condition and delay the progress of rehabilitation. In clinical practice, the Erasmus GBS Respiratory Insufficiency Score (EGRIS) can be used to predict respiratory failure and patients' need for mechanical ventilation.[31,32]

The patient's immune system attacks the myelin sheath of its own nervous system, causing damage to the myelin sheath and interfering with nerve signal transduction. Hyporeflexia and injury of glossopharyngeal and vagus nerve occurred. Glossopharyngeal and vagus nerve injury manifested as water choking cough and dysphagia. Dysphagia in patients with GBS may also aggravate respiratory muscle weakness, leading to respiratory complications such as aspiration pneumonia, which in turn aggravates the disease.

In summary, this article innovatively uses eight ML methods to construct models that affect the prognosis of GBS. This study found that the combination of NLR, age, mechanical ventilation, hyporeflexia and abnormal glossopharyngeal vagus nerve has a good predictive value for predicting the short-term prognosis of GBS patients. In this study, the RF, XGBoost and LGBM models performed well in predicting the prognosis of patients. This study found that the use of ML

to build a model to predict the prognosis of GBS patients is effective. The prediction model is helpful to predict and prevent poor prognosis in clinical practice.

There are limitations in this study. First, the sample size is relatively small. We searched the Medical Information Mart for Intensive Care (MIMIC) and eICU databases, but these two well-known databases have no relevant data to verify the model constructed in this study, so we used repeated sampling (bootstrap=1000) to make up for this problem. Repeated sampling is a very effective method for model validation, ensuring the credibility of the model.[13] In addition, we collected data with a different time period of patient admission than in the training set as a validation set to validate the model. Secondly, the characteristics included are not extensive enough. This study did not analyze the titer of anti-ganglioside antibodies and electrophysiological indicators. These indicators may also have an impact on the prognosis of patients. This study is a single center retrospective study. In the future, large-scale prospective studies are needed to verify and further explore predictive indicators for the severity of GBS, so as to assist physicians to evaluate and improve the prognosis of patients according to their individual data.

Abbreviations

ML, machine learning; GBS, Guillain- Barré syndrome; LASSO, Least Absolute Shrinkage and Selection Operator; DT, Decision Trees; RF, Random Forest; XGBoost, Extreme Gradient Boosting; KNN, k-nearest Neighbour; NB, Naive Bayes; NN, Neural Network; LGBM, Light Gradient Boosting Machine; LR, Logistic Regression; SHAP, SHapley additive explanation; ssGSEA, Single sample gene set enrichment analysis; AUC, area under the curve; NLR, neutrophil/lymphocyte ratio; IVIG, intravenous immunoglobulin; PE, plasma exchange; HFGS, Hughes functional grading scale; ROC, receiver operation characteristic; CAR, C-reactive protein/albumin ratio; EGRIS, Erasmus GBS Respiratory Insufficiency Score; MIMIC, Medical Information Mart for Intensive Care.

Data Sharing Statement

The datasets generated and/or analysed during this study are available from the corresponding author on reasonable request.

Ethics Approval

This study has obtained ethical approval from the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (2021-KY-0926).

Consent to Participate

All procedures performed in studies involving human participants were in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Consent for Publication

Not applicable.

Funding

"Four New" teaching reform research and practice project of Zhengzhou University in 2022 [2022ZZUSX021].

Disclosure

The authors declare no conflicts of interest in this work.

References

1. Willison HJ, Jacobs BC, van Doorn PA. Guillain-Barré syndrome. *Lancet*. 2016;388(10045):717–727. doi:10.1016/s0140-6736(16)00339-1
2. Bersano A, Carpo M, Allaria S, et al. Long term disability and social status change after Guillain-Barré syndrome. *J Neurol*. 2006;253(2):214–218. doi:10.1007/s00415-005-0958-x

3. Ruts L, Drenthen J, Jongen JLM, et al. Pain in Guillain-Barre syndrome: a long-term follow-up study. *Neurology*. 2010;75(16):1439–1447. doi:10.1212/WNL.0b013e3181f88345

4. Walgaard C, Jacobs BC, Lingsma HF, et al. Second intravenous immunoglobulin dose in patients with Guillain-Barré syndrome with poor prognosis (SID-GBS): a double-blind, randomised, placebo-controlled trial. *Lancet Neurol*. 2021;20(4):275–283. doi:10.1016/s1474-4422(20)30494-4

5. Singal AG, Mukherjee A, Elmunzer JB, et al. Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *Am J Gastroenterol*. 2013;108(11):1723–1730. doi:10.1038/ajg.2013.332

6. Lee SW, Lee EH, Choi IC. An ensemble machine learning approach to predict postoperative mortality in older patients undergoing emergency surgery. *BMC Geriatr*. 2023;23(1):262. doi:10.1186/s12877-023-03969-0

7. Wang X, Zhang X, Li H, et al. Application of machine learning algorithm in prediction of lymph node metastasis in patients with intermediate and high-risk prostate cancer. *J Cancer Res Clin Oncol*. 2023;149(11):8759–8768. doi:10.1007/s00432-023-04816-w

8. Sabbagh A, Washington SL, Tilki D, et al. Development and external validation of a machine learning model for prediction of lymph node metastasis in patients with prostate cancer. *Eur Urol Oncol*. 2023;6(5):501–507. doi:10.1016/j.euo.2023.02.006

9. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2(10):749–760. doi:10.1038/s41551-018-0304-0

10. Wen P, Wang L, Liu H, et al. Risk factors for the severity of Guillain-Barré syndrome and predictors of short-term prognosis of severe Guillain-Barré syndrome. *Sci Rep*. 2021;11(1):11578. doi:10.1038/s41598-021-91132-3

11. Wu X, Li C, Zhang B, et al. Predictors for mechanical ventilation and short-term prognosis in patients with Guillain-Barré syndrome. *Critical Care*. 2015;19(1):310. doi:10.1186/s13054-015-1037-z

12. Hiraga A, Kuwabara S. Early prediction of prognosis in Guillain-Barré syndrome. *Lancet Neurol*. 2007;6(7):572–573. doi:10.1016/s1474-4422(07)70155-2

13. Steyerberg EW. Validation in prediction research: the waste by data splitting. *J Clin Epidemiol*. 2018;103:131–133. doi:10.1016/j.jclinepi.2018.07.010

14. Bland JM, Altman DG. Statistics notes: bootstrap resampling methods. *BMJ*. 2015;350(jun02 13):h2622. doi:10.1136/bmj.h2622

15. Dwivedi AK, Mallawaarachchi I, Alvarado LA. Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method. *Stat Med*. 2017;36(14):2187–2205. doi:10.1002/sim.7263

16. Lee S, Ahmad A, Jeon G. Combining bootstrap aggregation with support vector regression for small blood pressure measurement. *J Med Sys*. 2018;42(4):63. doi:10.1007/s10916-018-0913-x

17. Auezova R, Ryskeldiev N, Doskaliyev A, et al. Association of preoperative levels of selected blood inflammatory markers with prognosis in gliomas. *Onco Targets Ther*. 2016;9:6111–6117. doi:10.2147/ott.S113606

18. Qun S, Tang Y, Sun J, et al. Neutrophil-to-lymphocyte ratio predicts 3-month outcome of acute ischemic stroke. *Neurotox Res*. 2017;31(3):444–452. doi:10.1007/s12640-017-9707-z

19. Imtiaz F, Shafique K, Mirza S, et al. Neutrophil lymphocyte ratio as a measure of systemic inflammation in prevalent chronic diseases in Asian population. *Int Arch Med*. 2012;5(1):2. doi:10.1186/1755-7682-5-2

20. de Jager CP, Wever PC, Gemen EFA, et al. The neutrophil-lymphocyte count ratio in patients with community-acquired pneumonia. *PLoS One*. 2012;7(10):e46561. doi:10.1371/journal.pone.0046561

21. Feng DY, Zhou Y-Q, Zhou M, et al. Risk factors for mortality due to ventilator-associated pneumonia in a Chinese hospital: a retrospective study. *Med Sci Monit*. 2019;25:7660–7665. doi:10.12659/msm.916356

22. Rajabally YA, Uncini A. Outcome and its predictors in Guillain-Barre syndrome. *J Neurol Neurosurg*. 2012;83(7):711–718. doi:10.1136/jnnp-2011-301882

23. Sipilä JO, Kauko T, Soilu-Hänninen M. Admission sodium level and prognosis in adult Guillain-Barré syndrome. *Int J Neurosci*. 2017;127(4):344–349. doi:10.3109/00207454.2016.1163551

24. Tunç A. Early predictors of functional disability in Guillain-Barré Syndrome. *Acta neurologica Belgica*. 2019;119(4):555–559. doi:10.1007/s13760-019-01133-3

25. Netto AB, Taly AB, Kulkarni GB, Uma Maheshwara Rao GS, Rao S. Prognosis of patients with Guillain-Barré syndrome requiring mechanical ventilation. *Neurol India*. 2011;59(5):707–711. doi:10.4103/0028-3886.86545

26. Ruiz-Sandoval JL, Salvatella-Gutiérrez A, López-Valencia G, et al. Clinical Characteristics and predictors of short-term outcome in Mexican adult patients with Guillain-Barré syndrome. *Neurol India*. 2021;69(1):107–114. doi:10.4103/0028-3886.310063

27. Minamino M, Miyamoto K, Kuwahara M, et al. Characteristics of Guillain-Barré syndrome in super-elderly individuals. *J Neurol*. 2023;270(4):2191–2196. doi:10.1007/s00415-023-11567-8

28. Liu S, Xiao Z, Lou M, et al. Guillain-Barré syndrome in southern China: retrospective analysis of hospitalised patients from 14 provinces in the area south of the Huaihe River. *J Neurol Neurosurg*. 2018;89(6):618–626. doi:10.1136/jnnp-2017-316930

29. Petzold A, Brettschneider J, Jin K, et al. CSF protein biomarkers for proximal axonal damage improve prognostic accuracy in the acute phase of Guillain-Barré syndrome. *Muscle and Nerve*. 2009;40(1):42–49. doi:10.1002/mus.21239

30. Zhang Y, Zhao Y, Wang Y. Prognostic factors of Guillain-Barré syndrome: a 111-case retrospective review. *Chin Neurosurg J*. 2018;4(1):14. doi:10.1186/s41016-018-0122-y

31. Walgaard C, Lingsma HF, Ruts L, et al. Prediction of respiratory insufficiency in Guillain-Barré syndrome. *Ann Neurol*. 2010;67(6):781–787. doi:10.1002/ana.21976

32. Yao J, Zhou R, Liu Y, et al. Predicting of mechanical ventilation and outcomes by using models and biomarker in Guillain-Barré syndrome. *Neurol Ther*. 2023;12(6):2121–2132. doi:10.1007/s40120-023-00546-w