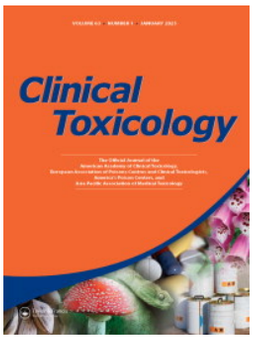

Clinical Toxicology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ictx20

Predicting delayed neurological sequelae in patients with carbon monoxide poisoning using machine learning models

Yunfeng Zhu, Tianshu Mei, Dawei Xu, Wei Lu, Dan Weng & Fei He

To cite this article: Yunfeng Zhu, Tianshu Mei, Dawei Xu, Wei Lu, Dan Weng & Fei He (14 Jan 2025): Predicting delayed neurological sequelae in patients with carbon monoxide poisoning using machine learning models, Clinical Toxicology, DOI: 10.1080/15563650.2024.2437113

To link to this article: https://doi.org/10.1080/15563650.2024.2437113

CLINICAL TOXICOLOGY
https://doi.org/10.1080/15563650.2024.2437113

Taylor & Francis
Taylor & Francis Group

CLINICAL RESEARCH

Predicting delayed neurological sequelae in patients with carbon monoxide poisoning using machine learning models

Yunfeng Zhu[a*], Tianshu Mei[b*], Dawei Xu[c*], Wei Lu[c], Dan Weng[a] and Fei He[b]

[a]School of Environmental and Biological Engineering, Nanjing University of Science & Technology, Nanjing, China; [b]Department of Emergency Medicine, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, China; [c]Department of Emergency Medicine, the Affiliated Suqian Hospital of Xuzhou Medical University, Suqian, China

ABSTRACT

Introduction: Delayed neurological sequelae is a common complication following carbon monoxide poisoning, which significantly affects the quality of life of patients with the condition. We aimed to develop a machine learning-based prediction model to predict the frequency of delayed neurological sequelae in patients with carbon monoxide poisoning.

Methods: A single-center retrospective analysis was conducted in an emergency department from January 01, 2018, to December 31, 2023. We analyzed data from patients with carbon monoxide poisoning, which were divided into training and test sets. We developed and evaluated sixteen machine learning models, using accuracy, sensitivity, specificity, and other relevant metrics. Threshold adjustments were performed to determine the most accurate model for predicting patients with carbon monoxide poisoning at risk of delayed neurological sequelae.

Results: A total of 360 patients with carbon monoxide poisoning were investigated in the present study, of whom 103 (28.6%) were diagnosed with delayed neurological sequelae, and two (0.6%) died. After threshold adjustment, the synthetic minority oversampling technique-random forest model demonstrated superior performance with an area under the receiver operating characteristic curve of 0.89 and an accuracy of 0.83. The sensitivity and specificity of the model were 0.9 and 0.8, respectively.

Discussion: The study developed a machine learning-based synthetic minority oversampling technique-random forest model to predict delayed neurological sequelae in patients with carbon monoxide poisoning, achieving an area under the receiver operating characteristic curve of 0.89. This technique was used to handle class imbalance, and shapley additive explanations analysis helped explain the model predictions, highlighting important factors such as the Glasgow Coma Scale, hyperbaric oxygen therapy, kidney function, immune response, liver function, and blood clotting.

Conclusions: The machine learning-based synthetic minority oversampling technique-random forest model developed in this study effectively identifies patients with carbon monoxide poisoning at high risk for delayed neurological sequelae.

ARTICLE HISTORY
Received 14 September 2024
Revised 20 November 2024
Accepted 27 November 2024

KEYWORDS
Carbon monoxide poisoning; delayed neurological sequelae; machine learning models; neurological deficits; prediction models

Introduction

Carbon monoxide poisoning is a major public health issue that can cause severe neurological damage or death [1]. In the United States (US), carbon monoxide poisoning accounts for approximately 50,000–100,000 emergency department visits and 1,500–2,000 fatalities each year [2]. A significant number of carbon monoxide poisoning survivors experience delayed neurological sequelae, a complication characterized by neurological, cognitive, and neuropsychiatric impairments resulting from brain injury [3]. Symptoms of delayed neurological sequelae range from mild personality changes to severe cognitive disturbances, emotional instability, language impairments, and focal neurological deficits, emerging days to weeks after acute carbon monoxide exposure [3,4]. Delayed neurological sequelae profoundly affects the quality of life of the patients and their families. Currently, several clinical indicators and imaging modalities, such as serum neuron-specific enolase activity [5], serum netrin-1 concentrations [6], and cranial diffusion-weighted magnetic resonance imaging [7], are being investigated for predicting delayed neurological sequelae in patients with carbon monoxide poisoning. However, the clinical utility of most biomarkers remains limited, as they are not widely available, and single indicators often show variability in performance across patients, making it difficult to fully reflect the complex nature of delayed neurological sequelae.

Machine learning models, which are mathematical or computational programs designed to identify patterns and make predictions on unseen datasets, offer a promising alternative [8]. Machine learning in healthcare depends on patient data collection; by organizing these data with specialized tools, algorithms

CONTACT Fei He hefei1201@njglyy.com Department of Emergency Medicine, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing 210008, China; Dan Weng danweng@njust.edu.cn School of Environmental and Biological Engineering, Nanjing University of Science & Technology, 200 Xiaolingwei, Nanjing, China.
*These authors contributed equally to the manuscript.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/15563650.2024.2437113.

in machine learning can reveal patterns that enhance diagnostic accuracy, develop personalized treatment plans, and improve outcome predictions for medical professionals. Compared with traditional statistical methods, machine learning models are constructed by the algorithm based on the data provided, without assuming a specific model structure. This flexibility allows machine learning to make fewer assumptions about data distribution, making it highly effective for analyzing diverse, unstructured, or high-dimensional datasets. Consequently, machine learning can account for numerous factors with complex relationships, leading to more accurate outcome predictions [9]. Furthermore, machine learning techniques can improve prediction models by detecting complex, multidimensional, non-linear patterns in the data [10]. Another key advantage is that a machine learning model can recognize trends automatically and continuously refine its performance over time [11]. These characteristics make machine learning particularly well-suited for tasks such as risk stratification, diagnosis and classification, and survival predictions in the medical field [12]. Emerging evidence suggests that machine learning models can improve the accuracy of prognosis predictions in poisoning-related conditions. A study by Veisani and colleagues [13] demonstrated that a machine learning-based model could identify risk factors for intentional and unintentional poisoning with an accuracy of 91.5%. Nevertheless, little is known about the role of machine learning in predicting delayed neurological sequelae frequency in patients with carbon monoxide poisoning. Hence, this study aims to develop an accurate machine learning model to identify patients at risk for delayed neurological sequelae, which can potentially enable early intervention.

Methods

Study design and setting

This single-center, retrospective study was conducted in the emergency department of Suqian Hospital of Nanjing Drum Tower Hospital group between January 1, 2018 and December 31, 2023. The hospital, located in northern China, is a 1,500-bed tertiary medical facility with approximately 210,000 annual emergency department visits, serving as a primary acute poisoning treatment center for both urban and rural areas. All adult patients with carbon monoxide poisoning were diagnosed based on the following criteria [14]: (1) a clinical history of carbon monoxide exposure; (2) the presence of any clinical signs or symptoms suggestive of carbon monoxide poisoning; and (3) an elevated carboxyhemoglobin level (>3% for non-smokers and >5% for smokers or those with unclear smoking status). Patients who were less than 18 years old, transferred from other hospitals, had incomplete clinical data, or were lost to follow-up were excluded from the study (Figure 1). We previously reported on patients with carbon monoxide poisoning hospitalized in our institution, from January 1, 2018, to December 31, 2020 [15].

Delayed neurological sequelae is clinically diagnosed based on the onset of neuropsychiatric symptoms, such as cognitive dysfunction, memory impairment, movement disorders, or focal neurological deficits, occurring within six weeks after the acute recovery phase of carbon monoxide poisoning [3]. Physicians routinely assessed the neurological deficits of the patients during hospitalization. For patients who were discharged within six weeks, follow-up telephone interviews were conducted to evaluate for delayed neurological sequelae. Based on these criteria, patients were categorized into two groups: those who developed delayed neurological sequelae (delayed neurological sequelae group) and those who did not (non-delayed neurological sequelae group).

To develop a machine learning model that could predict cases accurately and reliably, we employed the following strategies (Figure 2): (1) fine-grained preprocessing of the dataset, (2) the application of various machine learning techniques to enhance the model performance, (3) comparisons of different machine learning algorithms, (4) performance evaluation using multiple metrics, and (5) interpretation of the model output.

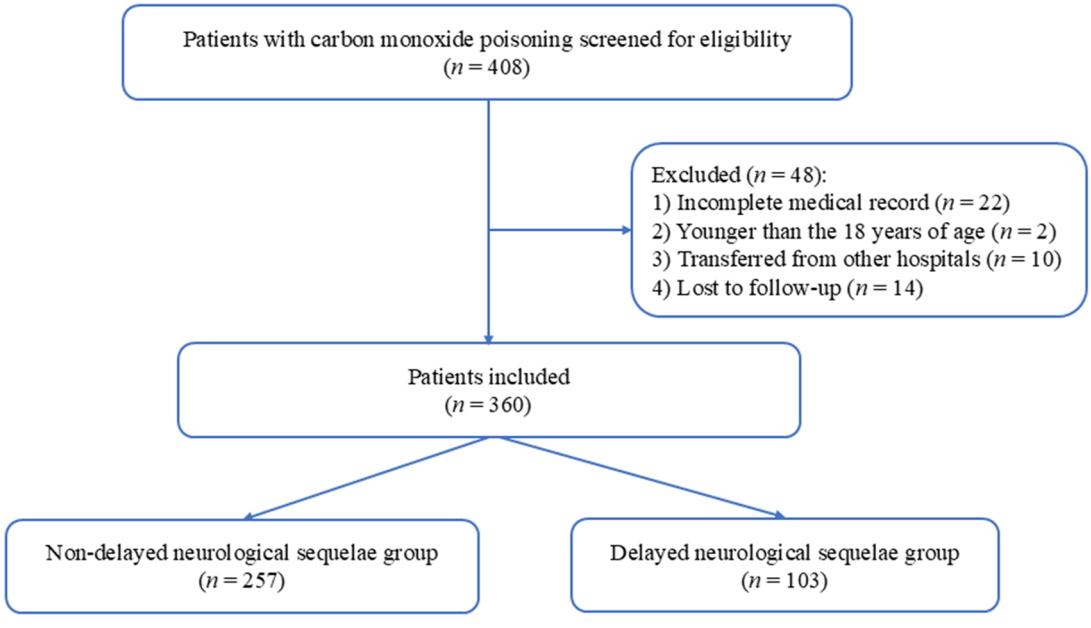

Figure 1. Patient selection and classification flow chart.

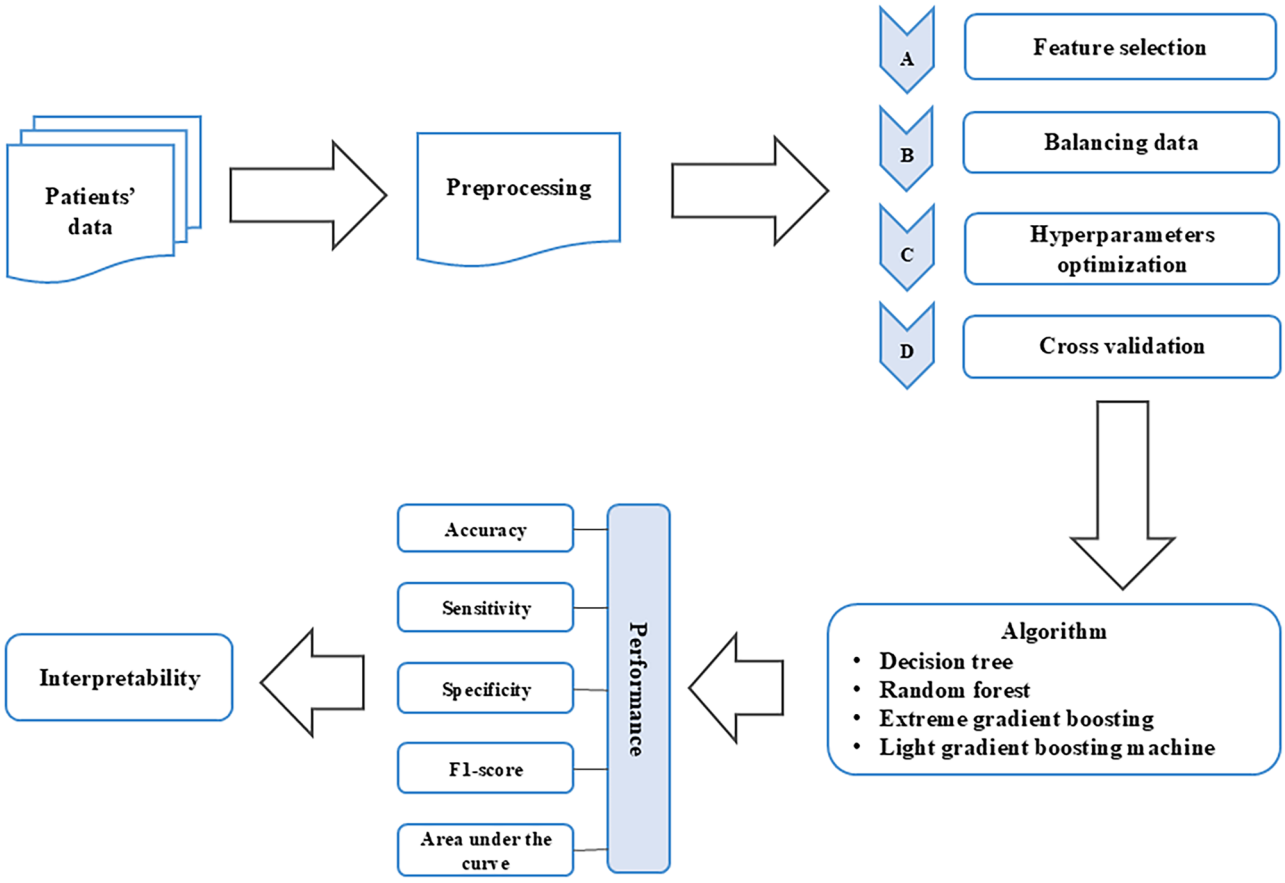

Figure 2. Roadmap of the proposed framework for predicting the frequency of delayed neurological sequelae in patients with carbon monoxide poisoning.

Ethics approval and consent to participate

This study was conducted in accordance with the Helsinki Declaration and Good Clinical Practice standards. It was approved by the Human Research Ethics Committee of the Affiliated Suqian Hospital of Xuzhou Medical University (EC 2021-025). The requirement for written informed consent was waived due to the retrospective and anonymized nature of the study.

Data collection

Demographic data, clinical characteristics, and laboratory parameters were collected at the time of admission and during hospitalization. Data collection was performed retrospectively by trained investigators using the electronic medical record system and subsequently reviewed by the research staff. These collected data included: age, gender, co-morbidities (hypertension, diabetes mellitus, coronary artery disease, and chronic obstructive pulmonary disease), Acute Physiology and Chronic Health Evaluation II (APACHE II) scores, Glasgow Coma Scale (GCS), heart rate, respiration rate, systolic blood pressure, diastolic blood pressure, mean arterial pressure, temperature, white blood cell count, neutrophil percentage, lymphocyte percentage, neutrophil-to-lymphocyte ratio (calculated by dividing the neutrophil count by the lymphocyte count), hemoglobin concentration, hematocrit, platelet count, red blood cell distribution width, prothrombin time, international normalized ratio, activated partial thromboplastin time, alanine aminotransferase activity, aspartate aminotransferase activity, total bilirubin concentration, direct bilirubin concentration, albumin concentration, globulin concentration, serum creatinine concentration, blood urea nitrogen concentration, uric acid concentration, estimated glomerular filtration rate (eGFR), pH, $PaCO_2$, PaO_2, $PaCO_2$/FiO2 ratio, serum lactate concentration, and blood glucose concentration). Interventions during hospitalization, including high-flow nasal cannula therapy, invasive mechanical ventilation, and hyperbaric oxygen therapy, were also recorded.

Software and packages used

The toolkits used for machine learning modeling and visualization were derived from Python 3.12.2 (https://www.Python.org/) and R 4.4.1 (https://www.R-project.org/). The Python libraries utilized included Pandas 2.2.2 (https://pandas.pydata.org/), Numpy 1.26.4 (https://numpy.org/), Matplotlib 3.9.0 (https://matplotlib.org/), seaborn 0.13.2 (https://seaborn.pydata.org/), pillow 10.3.0 (https://python-pillow.org/), imbalanced-learn 0.12.3 (https://imbalanced-learn.org/stable/), scikit-learn 1.5.0 (https://scikit-learn.org/), XGBoost 2.1.0 (https://xgboost.readthedocs.io/en/stable/), Lightgbm 4.4.0 (https://lightgbm.readthedocs.io/en/latest/), Bayesian-optimization 1.4.3 (https://bayesian-optimization.github.io/BayesianOptimization/index.html) and SHAP (SHapley Additive exPlanations) 0.45.1 (https://shap.readthedocs.io/en/latest/) for data analysis, modeling, and visualization. In addition, five R

packages were employed for data analysis and visualization: Tidyverse 2.0.0 (https://tidyverse.tidyverse.org/), Psych 2.4.6.26 (https://CRAN.R-project.org/package=psych), RColorBrewer 1.1-3 (https://CRAN.R-project.org/package=RColorBrewer), Corrplot 0.92 (https://github.com/taiyun/corrplot), and Ggstatsplot 0.12.3 (https://indrajeetpatil.github.io/ggstatsplot/).

Preprocessing

We split the dataset into a training set and a test set in the ratio of 8:2 and ensured that the proportion of different categories in training and test sets remained consistent with the initial dataset. We then standardized the numerical-type features in training and test sets so that they were distributed as a standard normal distribution with mean of zero and a standard deviation of 1.

Feature selection

We applied logistic regression with L1 regularization to identify the 10 most important features. L1 regularization adds the L1 norm of feature weights to the loss function, encouraging certain feature weights to shrink towards zero [16].

Class imbalance

Our dataset exhibited class imbalance, with significantly fewer patients developing delayed neurological sequelae compared to those who did not. To address this, we applied three oversampling techniques: synthetic minority oversampling technique, adaptive synthetic sampling, and random oversampling to balance the training set. Additionally, we trained models using the original unbalanced dataset for comparison with models trained on the balanced set.

Model selection and hyperparameter optimization

We selected four tree-based models to develop our delayed neurological sequelae prediction models: decision tree, random forest, extreme gradient boosting, and light gradient boosting machine. These models were combined with four resampling strategies, resulting in a total of 16 models. For clarity, these models were named as follows, no treatment-decision tree, no treatment-random forest, no treatment-extreme gradient boosting, no treatment-light gradient boosting machine, synthetic minority oversampling technique-decision tree, synthetic minority oversampling technique-random forest, synthetic minority oversampling technique-extreme gradient boosting, synthetic minority oversampling technique-light gradient boosting machine, adaptive synthetic sampling-decision tree, adaptive synthetic sampling-random forest, adaptive synthetic sampling-extreme gradient boosting, adaptive synthetic sampling-light gradient boosting machine, random oversampling-decision tree, random oversampling-random forest, random oversampling-extreme gradient boosting, and random oversampling-light gradient boosting machine. We used 10-fold cross-validation combined with Bayesian optimization to tune the hyperparameters of all models.

Model evaluation and explanation

We selected several metrics to comprehensively evaluate our models, including confusion matrix, accuracy, specificity, sensitivity, precision and F1-score. The confusion matrix compared the predicted results of the model with the actual class labels (Table 1), and other evaluation metrics were derived from it (Table 2). Additionally, we used receiver operating characteristic curves to visualize the performance of all models. To further interpret the outputs of our tree-based models, we employed shapley additive explanations, a robust tool for explaining and visualize machine learning model outputs [17].

Statistical analyses

All statistical analyses were conducted using SPSS version 22.0 for Windows (SPSS Inc., Chicago, IL, USA). Continuous variables with a normal distribution were presented as mean ± standard deviation and compared using the Student's t-test. Non-normally distributed continuous data were reported as medians with interquartile ranges (IQR) and analyzed using the Mann-Whitney U test. Categorical variables were expressed as numbers and percentages and compared using either the chi-square test or Fisher's exact test, as appropriate. Two-tailed $P<0.05$ was considered statistically significant.

Results

Patient characteristics

A total of 408 patients diagnosed with carbon monoxide poisoning were eligible for the study, with 360 patients meeting the inclusion criteria (Figures 1 and 3A). The mean age of the patients was 54 ± 18 years and 58.6% (211/360) were female. The overall frequency of delayed neurological sequelae was 28.6% (103/360) with a 28-day mortality rate

Table 1. Confusion matrix for classification results.

	Predicted value	
True value	Positive	Negative
Positive	True positive	False negative
Negative	False positive	True negative

Rows represent the actual categories, while columns represent the predicted categories.

Table 2. Metrics to measure model performance.

Metrics	Formula
Accuracy	$\frac{True\ positive + True\ negative}{True\ positive + False\ positive + False\ negative + True\ negative}$
Specificity	$\frac{True\ negative}{True\ negative + False\ positive}$
Precision	$\frac{True\ positive}{True\ positive + False\ positive}$
Recall (sensitivity)	$\frac{True\ positive}{True\ positive + False\ negative}$
F1-score	$\frac{2 \times precision \times recall}{precision + recall}$

of 0.6% (2/360). A comparison of baseline clinical features between the delayed neurological sequelae and non-delayed neurological sequelae groups is presented in Table 3. Patients in the delayed neurological sequelae group had significantly higher values for age, APACHE II scores, heart rate, respiratory rate, white blood cell count, neutrophil percentage, neutrophil-to-lymphocyte ratio, red blood cell distribution width, prothrombin time, activated partial thromboplastin time, alanine aminotransferase activity, aspartate aminotransferase activity, serum creatinine concentration, blood urea nitrogen concentration, serum lactate concentration, and blood glucose concentration compared to those in the non-delayed neurological sequelae group. Additionally, the proportion of patients requiring invasive mechanical ventilation was markedly higher in the delayed neurological sequelae group. Conversely, patients in the delayed neurological sequelae group had significantly lower values for the proportion of females, Glasgow Coma Scale, lymphocyte percentage, albumin concentration, globulin concentration, eGFR, pH, and PaO_2/FiO_2 ratio. There was no significant difference in the 28-day mortality or length of hospital stay between the two groups.

Table 3. Characteristics, management and outcomes of 360 patients with carbon monoxide poisoning with or without delayed neurological sequelae.

Variable	All ($n=360$)	Delayed neurological sequelae group ($n=103$)	Non-delayed neurological sequelae group ($n=257$)	P value
Age (years), mean ± SD	53.7 ± 19.0	59.5 ± 20.3	51.4 ± 17.9	<0.001
Female, *n* (%)	211 (58.6)	44 (42.7)	167 (65.0)	<0.001
Comorbidities, *n* (%)				
Hypertension	72 (20.0)	27 (26.2)	45 (17.5)	0.06
Diabetes mellitus	17 (4.7)	8 (7.8)	9 (3.5)	0.09
Coronary artery disease	15 (4.1)	7 (6.8)	8 (3.1)	0.11
Chronic obstructive pulmonary disease	13 (3.6)	6 (5.8)	7 (2.7)	0.15
Glasgow Coma Scale, mean ± SD	9.6 ± 3.3	6.7 ± 2.3	10.7 ± 3.0	<0.001
APACHE II scores, median (IQR)	12.0 (8.0–17.0)	17.0 (14.0–21.0)	10.0 (7.0–15.0)	<0.001
Heart rate (beats/min), mean ± SD	75.6 ± 18.5	81.4 ± 21.7	73.2 ± 16.5	<0.001
Respiration rate (breaths/min), median (IQR)	18.0 (18.0–19.0)	18.0 (18.0–22.0)	18.0 (18.0–18.0)	0.01
Systolic blood pressure (mmHg), mean ± SD	126.5 ± 27.8	126.0 ± 30.1	126.8 ± 26.8	0.83
Diastolic blood pressure (mmHg), mean ± SD	77.4 ± 18.0	79.3 ± 19.9	76.6 ± 17.1	0.49
Mean arterial blood pressure (mmHg), median (IQR)	93.7 (74.1–106.6)	96.7 (70.3–111.7)	93.3 (75.3–104.5)	0.92
Temperature (°C), mean ± SD	36.7 ± 0.5	36.8 ± 0.7	36.7 ± 0.5	0.13
White blood cell count ($\times10^9$/L), mean ± SD	9.3 ± 4.0	10.9 ± 4.3	8.7 ± 3.7	<0.001
Neutrophil percentage (%), mean ± SD	71.5 ± 14.3	77.8 ± 12.5	69.1 ± 14.4	<0.001
Lymphocyte percentage (%), median (IQR)	18.4 (10.9–28.7)	13.0 (7.4–19.2)	21.5 (13.6–31.9)	<0.001
Neutrophil-to-lymphocyte ratio ($\times10^9$/L), median (IQR)	4.0 (2.2–7.4)	6.00 (3.8–11.7)	3.3 (1.9–6.0)	<0.001
Hemoglobin concentration (g/L), mean ± SD	132.1 ± 22.1	128.1 ± 28.4	133.6 ± 18.8	0.07
Platelets count ($\times10^9$/L), mean ± SD	213.3 ± 58.7	211.6 ± 60.8	214.0 ± 58.0	0.73
Hematocrit (%), mean ± SD	0.4 ± 0.1	0.4 ± 0.1	0.4 ± 0.1	0.15
Red blood cell distribution width (%), mean ± SD	12.9 ± 1.3	13.2 ± 1.4	12.9 ± 1.3	0.04
Prothrombin time (s), mean ± SD	10.8 ± 1.0	11.0 ± 0.9	10.8 ± 1.0	0.03
Activated partial thromboplastin time (s), mean ± SD	25.2 ± 3.6	25.9 ± 3.5	25.0 ± 3.6	0.02
International normalized ratio, mean ± SD	0.9 ± 0.1	1.0 ± 0.1	0.9 ± 0.1	0.16
Total bilirubin (μmol/L), median (IQR)	10.3 (8.0–14.0)	10.4 (8.0–14.0)	10.1 (8.0–14.0)	0.77
Total bilirubin concentration (mg/dL), median (IQR)	0.6 (0.5–0.8)	0.6 (0.5–0.8)	0.6 (0.5–0.8)	0.67
Direct bilirubin (μmol/L), mean ± SD	3.3 ± 2.1	3.5 ± 1.9	3.3 ± 2.3	0.27
Direct bilirubin concentration (mg/dL), mean ± SD	0.2 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.33
Alanine aminotransferase activity (U/L), median (IQR)	19.8 (14.0–27.1)	22.0 (16.2–37.9)	19.0 (13.5–26.0)	0.02
Aspartate aminotransferase activity (U/L), median (IQR)	23.0 (17.3–30.0)	28.0 (21.0–47.0)	21.6 (17.0–27.7)	<0.001
Albumin concentration (g/L), mean ± SD	40.7 ± 4.7	38.9 ± 4.6	41.4 ± 4.6	<0.001
Globulin concentration (g/L), mean ± SD	26.0 ± 5.6	25.1 ± 5.4	26.4 ± 5.6	0.04
Serum creatinine (μmol/L), mean ± SD	63.8 ± 18.8	76.8 ± 22.3	58.6 ± 14.4	<0.001
Serum creatinine concentration (mg/dL), mean ± SD	0.7 ± 0.2	0.9 ± 0.3	0.7 ± 0.2	<0.001
Blood urea nitrogen (mmol/L), mean ± SD	6.1 ± 2.2	7.0 ± 2.5	5.8 ± 2.0	<0.001
Blood urea nitrogen concentration (mg/dL), mean ± SD	17.2 ± 6.3	19.6 ± 7.0	16.2 ± 5.7	<0.001
Uric acid (μmol/L), mean ± SD	289.3 ± 107.6	301.7 ± 111.9	284.3 ± 105.6	0.17
Uric acid concentration (mg/dL), mean ± SD	4.9 ± 1.8	5.1 ± 1.9	4.8 ± 1.8	0.16
eGFR (mL/min/1.73m^2), mean ± SD	101.5 (86.3–119.4)	90.2 (78.3–102.2)	107.5 (90.6–123.3)	<0.001
Serum lactate concentration (mmol/L), median (IQR)	2.0 (1.5–2.9)	2.4 (1.6–3.7)	1.8 (1.4–2.6)	<0.001
pH (mean ± SD)	7.4 ± 0.1	7.4 ± 0.1	7.4 ± 0.1	0.01
PaO_2 (mmHg), median (IQR)	88.2 (74.0–103.1)	85.8 (72.4–105.9)	90.4 (76.1–103.0)	0.262
$PaCO_2$, (mmHg), mean ± SD	36.7 ± 6.0	36.9 ± 6.7	36.7 ± 5.7	0.76
PaO_2/FiO2 ratio (mmHg), mean ± SD	205.6 ± 53.4	191.8 ± 53.3	211.1 ± 52.5	0.02
Blood glucose (mmol/L), mean ± SD	7.2 ± 3.2	8.2 ± 4.6	6.8 ± 2.3	0.03
Blood glucose concentration (mg/dL), mean ± SD	130.0 ± 57.7	148.3 ± 83.4	122.6 ± 41.3	0.01
Interventions *n* (%)				
High-flow nasal cannula	304 (84.4)	89 (86.4)	215 (83.7)	0.52
Invasive mechanical ventilation	8 (2.2)	8 (7.8)	0 (0)	<0.001
Hyperbaric oxygen therapy	229 (63.6)	63 (61.2)	166 (64.6)	0.54
Outcomes				
28-day mortality *n* (%)	2 (0.5)	2 (1.9)	0 (0)	0.08
Hospital stays (days), median (IQR)	6 (4–9)	6 (3–9)	6 (5–9)	0.13

APACHE II, acute physiology and chronic health evaluation II; eGFR, estimated glomerular filtration rate; IQR, interquartile range; SD, standard deviation

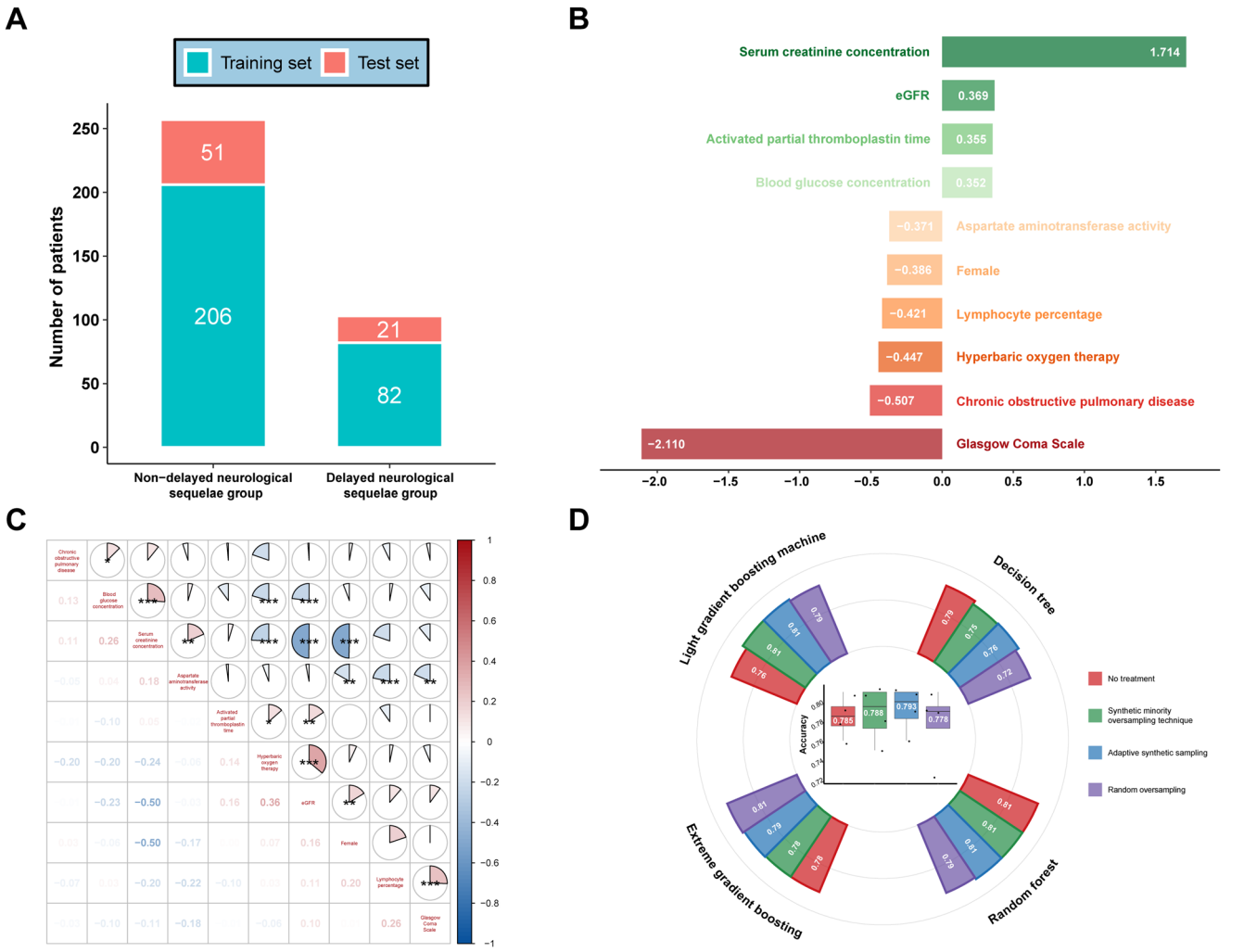

Figure 3. Data distribution, features selected, heat map of correlation between features selected and the accuracy of 16 models in the test set. (A) The number of carbon monoxide patients with or without delayed neurological sequelae in the training set and test set. (B) The significance of the 10 selected demographic, clinical and intervention features. The length of the bars and the absolute value of the coefficients reflect the importance of each feature. (C) The strength of correlations between the 10 selected demographic, clinical and intervention features. The size of the pie charts, magnitude of values, and intensity of colors indicate the degree of correlation. Significance levels are marked with symbols: * $P<0.05$; ** $P<0.005$; *** $P<0.001$. (D) The accuracy of the 16 models in the test set (the outer circle) and the average accuracy of the models using each of the four balancing methods (box plot in the center).

Feature selection

Logistic regression not only identified the 10 most important features, but also provided coefficients for these features (Figure 3B). Additionally, we observed no strong correlation among them (Figure 3C). This indicated that the selected features each contribute unique information, which was crucial for constructing a robust model.

Model selection

We optimized the hyperparameters of all models after balancing the training set. The optimal hyperparameters for each model are detailed in Supplemental Table S1. We then conducted a comprehensive performance assessment of all models. As shown in Supplemental Figure S1, all 16 models demonstrated commendable accuracy on the training set, and the choice of balancing strategy did not significantly impact their performance. However, notable discrepancies emerged in the performance when evaluating the test set (Figures 3D and 4, Table 3). Among the models, synthetic minority oversampling technique-decision tree, synthetic minority oversampling technique-random forest, and adaptive synthetic sampling-random forest models exhibited superior predictive capabilities for identifying patients at risk of delayed neurological sequelae. Receiver operating characteristic curves for these three models are plotted in Figure 5A, followed by a detailed analysis.

Most classifiers use a default threshold of 0.5 for predictive probability. A sample with a probability below 0.5 is classified as negative, while a probability above 0.5 indicates a positive classification. We adjusted the threshold for these three models to improve prediction accuracy for delayed neurological sequelae. To identify the optimal threshold, we varied the threshold between 0 and 0.5 and used the F1-score to identify the optimal threshold. Figure 5B shows the F1-score for the three models across different thresholds. The synthetic minority oversampling technique-random forest

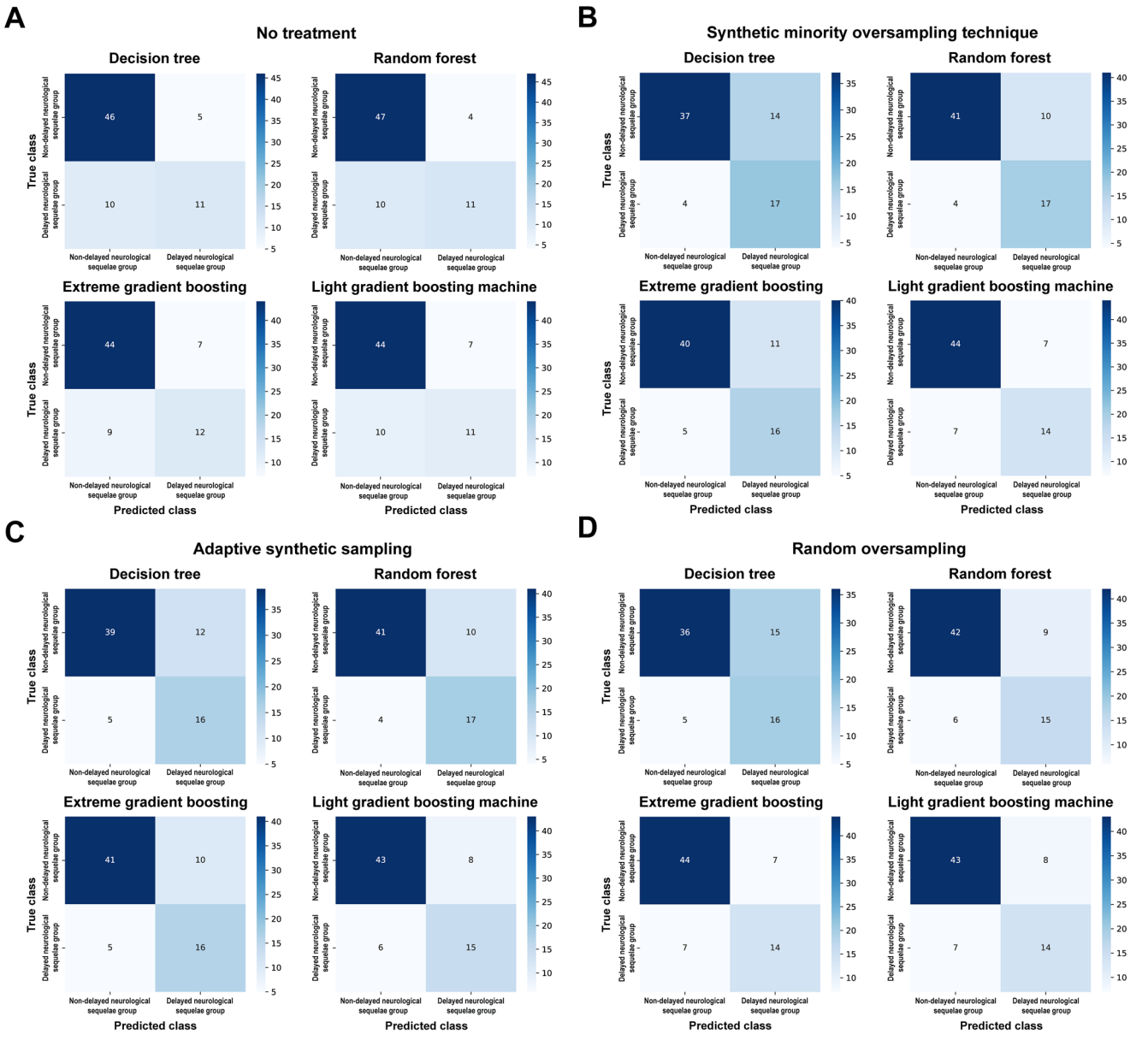

Figure 4. Confusion matrices of 16 machine learning models.

model achieved optimal performance with a threshold set to approximately 0.485.

After threshold adjustment, the synthetic minority oversampling technique-random forest model achieved an accuracy of 0.83, sensitivity of 0.9, and specificity of 0.8, reflecting significant improvements over its pre-adjustment performance. The confusion matrix for this model is presented in Figure 5C. Consequently, we selected the synthetic minority oversampling technique-random forest model as our final model.

Model explanation

The random forest classifier, which comprises numerous decision trees, can be complex to interpret. To address this challenge, we used shapley additive explanations to clarify the prediction mechanism of the synthetic minority oversampling technique-random forest model. As shown in Figure 5D, high values for six specific predictors (GCS, eGFR, lymphocyte percentage, hyperbaric oxygen therapy, female, and chronic obstructive pulmonary disease) were associated with negative shapley additive explanations values, indicating that higher value for these features was linked to lower likelihood of developing delayed neurological sequelae. In contrast, features such as serum creatinine concentration, aspartate aminotransferase activity, blood glucose concentration, and activated partial thromboplastin time exhibited positive shapley additive explanations values at higher values, suggesting that elevated values for these predictors were associated with increased risk of delayed neurological sequelae.

Discussion

Early prediction of neurological prognosis in patients with carbon monoxide poisoning can help identify high-risk individuals, allowing timely interventions and better-informed care decisions, potentially improving long-term outcomes and reducing

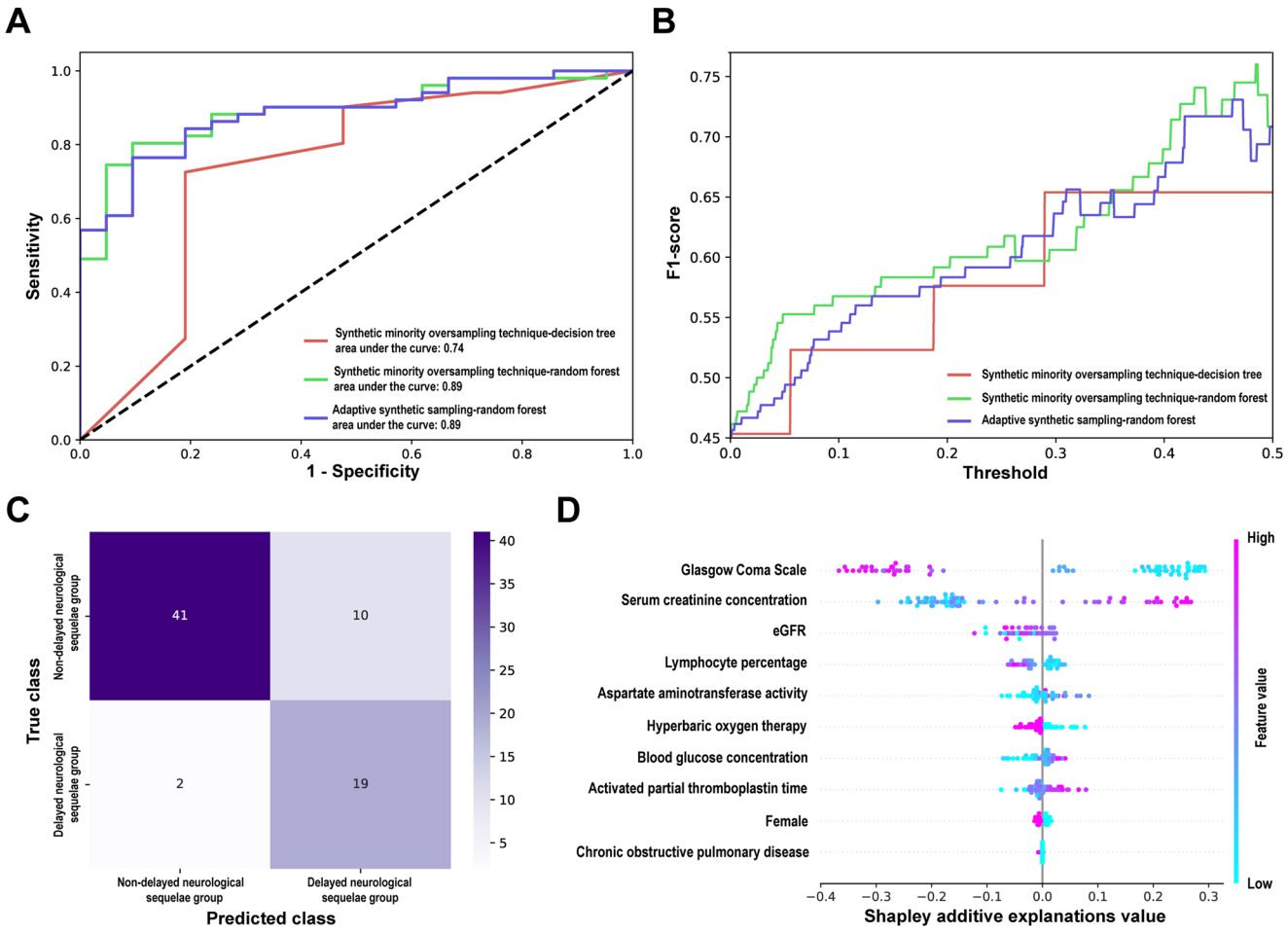

Figure 5. Performance evaluation and feature importance analysis of the synthetic minority oversampling technique-random Forest model. (A) Receiver operating characteristic curves for three machine learning models (synthetic minority oversampling technique-decision tree, synthetic minority oversampling technique-random Forest, and adaptive synthetic sampling-random Forest) tested on the same cohort. (B) Threshold adjustment curve for three machine learning models (synthetic minority oversampling technique-decision tree, synthetic minority oversampling technique-random Forest, and adaptive synthetic sampling-random Forest), illustrating the change in F1-score across 1,000 evenly spaced threshold values between 0 and 0.5. (C) Confusion matrix of the synthetic minority oversampling technique-random Forest model after threshold adjustment. (D) Shapley additive explanations bee swarm plot showing feature importance for the synthetic minority oversampling technique-random Forest model. Features are ranked by the average absolute shapley additive explanations value. Each point represents a sample, with jitter added for dispersion where samples share the same value. The x-axis represents shapley additive explanations values, with positive values indicating a prediction of delayed neurological sequelae and negative values indicating a prediction of non-delayed neurological sequelae.

healthcare costs. In the present study, we developed a machine learning-based model to predict delayed neurological sequelae in patients with carbon monoxide poisoning. Our results indicated that the synthetic minority oversampling technique-random forest model demonstrated robust predictive value for predicting delayed neurological sequelae, with an area under the receiver operating characteristic curve of 0.89.

Synthetic minority oversampling technique, an advanced oversampling technique, performs interpolation among neighboring minority class instances to address class imbalance [18,19]. By applying the synthetic minority oversampling technique to the dataset, we reduced bias toward the majority class, thus enhancing performance for the minority class. Random forest, a popular ensemble learning algorithm, builds a "forest" of decision trees and aggregates their results to produce a final prediction. Known for its ease of implementation and low computational cost, random forest consistently delivers strong performance in real-world applications and is recognized as a leading method in ensemble learning [20].

Researchers have investigated various clinical indicators to assess the risk of delayed neurological sequelae in patients with carbon monoxide poisoning at an early stage, such as serum lactate concentrations [21] and QT interval prolongation [22]. However, these indicators demonstrated significantly lower area under the receiver operating characteristic curve values, compared to the machine learning-based approach employed in the present study. Consistent with previous reports, our study demonstrated that machine learning models can rapidly analyze extensive data from the electronic health records of patients with carbon monoxide poisoning, including demographic and clinical characteristics, to predict which patients are at high risk of developing delayed neurological sequelae. This may allow for timely treatment and management recommendations. Furthermore, machine learning models can automatically identify underlying patterns in data, which may align with clinical judgment or reveal previously unknown patterns.

In the present study, we used shapley additive explanations to elucidate the internal mechanisms of the synthetic minority

oversampling technique-random forest model (Figure 5D). This analysis underscored the importance of the Glasgow Coma Scale and hyperbaric oxygen therapy in predicting delayed neurological sequelae, aligning with findings from other studies [23,24]. It is worth noting that the proportion of patients receiving hyperbaric oxygen therapy was similar between the delayed neurological sequelae and non-delayed neurological sequelae groups in our study, suggesting that hyperbaric oxygen therapy alone may not be a significant predictor of delayed neurological sequelae. This highlights the machine learning models likely rely on a broader range of variables, not just hyperbaric oxygen therapy, for accurate predictions. Additionally, the model identified that kidney function as an important factor in delayed neurological sequelae development, with higher serum creatinine concentration and lower eGFR increasing the likelihood of delayed neurological sequelae. The results also revealed associations between lower lymphocyte percentage, higher alanine aminotransferase activity and blood glucose concentration, and prolonged activated partial thromboplastin time with a higher risk of delayed neurological sequelae in patients with carbon monoxide poisoning. These findings suggest that delayed neurological sequelae are associated with multiple functional impairments, including immune system dysfunction, liver function abnormalities, glucose abnormalities, and endogenous coagulation disturbances. In addition, our study indicated a potential gender difference in delayed neurological sequelae risk, with female patients with carbon monoxide poisoning exhibiting a lower frequency of delayed neurological sequelae, possibly due to higher hypoxia tolerance [25].

The present study has several limitations that must be acknowledged. First, this is a single-center study with a relatively small sample size and no external validation. Therefore, future works with larger sample sizes and across multiple centres are needed to further assess the generalizability of the model across different populations and clinical settings. Second, there is no universally accepted set of diagnostic criteria for delayed neurological sequelae following carbon monoxide poisoning. Therefore, the variations in how delayed neurological sequelae is defined may affect the generalizability of our results. Future research is needed to help standardize definitions and enhance comparability across studies. Third, previous study has suggested that patients with persistent neurological symptoms after carbon monoxide poisoning may be at higher risk of developing delayed neurological sequelae [26]. In the present study, we did not distinguish between persistent neurological symptoms and delayed neurological sequelae in the included population. This lack of differentiation may affect the interpretation of our predictive model for delayed neurological sequelae. Future analyses that separate persistent neurological symptoms and delayed neurological sequelae may lead to changes in the predictive performance of the model and improve the understanding of factors specifically contributing to delayed neurological sequelae development. Finally, we did not include clinical parameters that were missing in more than half of the patients, such as troponin concentration, duration of carbon monoxide exposure, smoking status, and interval from exposure termination to emergency department arrival since such missing data could impact the machine learning model in several ways, such as reduced model performance and algorithm instability

In conclusion, our study demonstrated that the synthetic minority oversampling technique-random forest model can accurately predict delayed neurological sequelae in patients with carbon monoxide poisoning. This model could assist clinical toxicologists in identifying patients at risk for personalized management plans.

Acknowledgement

We would like to thank Ms. Xiaoshun Fang for giving us so much encouragement and support during this research period.

Authors contributions

Dan Weng and Fei He contributed to the conception and design of the study; Yunfeng Zhu, Tianshu Mei and Dawei Xu performed research; Yunfeng Zhu, Wei Lu and Dawei Xu performed analysis and interpretation of data; Yunfeng Zhu and Fei He contributed to drafting the article or revising it critically for important intellectual content; All authors final approval of the version to be submitted.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by Jiangsu Traditional Chinese Medicine Science and Technology Development Plan Project, China (MS2023060), Medical Science and technology development Foundation, Nanjing Department of Health (YKK22068) and Jiangsu Health International Exchange Program.

ORCID

Yunfeng Zhu http://orcid.org/0009-0009-4010-7973
Tianshu Mei http://orcid.org/0009-0001-9693-0776
Dawei Xu http://orcid.org/0000-0002-5877-8769
Dan Weng http://orcid.org/0000-0001-8078-6864
Fei He http://orcid.org/0000-0003-3721-198X

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

[1] Chenoweth JA, Albertson TE, Greer MR. Carbon monoxide poisoning. Crit Care Clin. 2021;37(3):657–672. doi: 10.1016/j.ccc.2021.03.01.
[2] Dent MR, Rose JJ, Tejero J, et al. Carbon monoxide poisoning: from microbes to therapeutics. Annu Rev Med. 2024;75(1):337–351. doi: 10.1146/annurev-med-052422-020045.
[3] Rose JJ, Wang L, Xu Q, et al. Carbon monoxide poisoning: pathogenesis, management, and future directions of therapy. Am J Respir Crit Care Med. 2017;195(5):596–606. doi: 10.1164/rccm.201606-1275CI.
[4] Choi S, Nah S, Han S. Correlation between time to hyperbaric oxygen therapy and delayed neurological sequelae in acute carbon monoxide poisoning patients. Diagnostics (Basel). 2024;14(2):186. doi: 10.3390/diagnostics14020186.

[5] Nah S, Choi S, Kim GW, et al. Prediction of delayed neuropsychiatric sequelae after carbon monoxide poisoning via serial determination of serum neuron-specific enolase levels. Hum Exp Toxicol. 2021;40(12_suppl):S339–S346. doi: 10.1177/09603271211043475.
[6] Kokulu K, Mutlu H, Sert ET. Serum netrin-1 levels at presentation and delayed neurological sequelae in unintentional carbon monoxide poisoning. Clin Toxicol (Phila). 2020;58(12):1313–1319. doi: 10.1080/15563650.2020.1743302.
[7] Gao X, Wei W, Yang GD. Clinical factors for delayed neuropsychiatric sequelae from acute carbon monoxide poisoning: a retrospective study. Front Med. 2024;11:1333197. doi: 10.3389/fmed.2024.1333197.
[8] Handelman GS, Kok HK, Chandra RV, et al. eDoctor: machine learning and the future of medicine. J Intern Med. 2018;284(6):603–619. doi: 10.1111/joim.12822.
[9] Ley C, Martin RK, Pareek A, et al. Machine learning and conventional statistics: making sense of the differences. Knee Surg Sports Traumatol Arthrosc. 2022;30(3):753–757. doi: 10.1007/s00167-022-06896-6.
[10] Deo RC. Machine learning in medicine: will this time be different? Circulation. 2020;142(16):1521–1523. doi: 10.1161/CIRCULATIONAHA.120.050583.
[11] Krishnan G, Singh S, Pathania M, et al. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. Front Artif Intell. 2023;6:1227091. doi: 10.3389/frai.2023.1227091.
[12] Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019;20(5):e262–e273. doi: 10.1016/S1470-2045(19)30149-4.
[13] Veisani Y, Sayyadi H, Sahebi A, et al. Comparison of machine learning algorithms to predict intentional and unintentional poisoning risk factors. Heliyon. 2023;9(6):e17337. doi: 10.1016/j.heliyon.2023.e17337.
[14] Kim YJ, Sohn CH, Seo DW, et al. Analysis of the development and progression of carbon monoxide poisoning-related acute kidney injury according to the kidney disease improving global outcomes (KDIGO) criteria. Clin Toxicol (Phila). 2018;56(8):759–764. doi: 10.1080/15563650.2018.1424890.
[15] Xu DW, Mei TS, He F. The neutrophil-to-lymphocyte ratio is associated with the frequency of delayed neurologic sequelae in patients with carbon monoxide poisoning. Sci Rep. 2023;13(1):19706. doi: 10.1038/s41598-023-47214-5.
[16] Zhou Z-H. Chapter 11, Feature selection and sparse learning. In: Machine learning. Singapore: Springer; 2021. p. 265–285. doi: 10.1007/978-981-15-1967-3_11.
[17] Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020;2(1):56–67. doi: 10.1038/s42256-019-0138-9.
[18] Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. jair. 2002;16:321–357. doi: 10.48550/arXiv.1106.1813.
[19] Fernandez A, Garcia S, Herrera F, et al. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. JAIR. 2018;61:863–905. doi: 10.1613/jair.1.11192.
[20] Rigatti SJ. Random forest. J Insur Med. 2017;47(1):31–39. doi: 10.17849/insm-47-01-31-39.1.
[21] Lee H, Oh J, Kang H, et al. Association between early phase serum lactate levels and occurrence of delayed neuropsychiatric sequelae in adult patients with acute carbon monoxide poisoning: a systematic review and meta-analysis. J Pers Med. 2022;12(4):651. doi: 10.3390/jpm12040651.
[22] Liao SC, Mao YC, Hung YM, et al. Predictive role of QTc prolongation in carbon monoxide poisoning-related delayed neuropsychiatric sequelae. Biomed Res Int. 2018;2018:2543018–2543018. doi: 10.1155/2018/2543018.
[23] Thom SR. Hyperbaric-oxygen therapy for acute carbon monoxide poisoning. N Engl J Med. 2002;347(14):1105–1106. doi: 10.1056/NEJMe020103.
[24] Xu P, Wang Y, Cao L, et al. Glasgow Coma scale is a better delayed neurological sequelae risk factor than neurological examination abnormalities in carbon monoxide poisoning. Am J Emerg Med. 2020;38(11):2468–2469. doi: 10.1016/j.ajem.2020.02.047.
[25] Botek M, Krejčí J, McKune A. Sex differences in autonomic cardiac control and oxygen saturation response to short-term normobaric hypoxia and following recovery: effect of aerobic fitness. Front Endocrinol. 2018;9:697. doi: 10.3389/fendo.2018.00697.
[26] Suzuki Y. Risk factors for delayed encephalopathy following carbon monoxide poisoning: importance of the period of inability to walk in the acute stage. PLOS One. 2021;16(3):e0249395. doi: 10.1371/journal.pone.0249395.