Open Access Full Text Article

ORIGINAL RESEARCH

Development and Validation of an ICU-Venous Thromboembolism Prediction Model Using Machine Learning Approaches: A Multicenter Study

Jie Jin [1], Jie Lu[1], Xinyang Su[2], Yinhuan Xiong[3], Shasha Ma[4], Yang Kong[5], Hongmei Xu[1]

[1]School of Nursing, Binzhou Medical University, Binzhou, People's Republic of China; [2]Department of Spine Surgery, Binzhou Medical University Hospital, Binzhou, People's Republic of China; [3]Department of Nursing, Binzhou People's Hospital, Binzhou, People's Republic of China; [4]Department of Neurosurgery, Binzhou Medical University Hospital, Binzhou, People's Republic of China; [5]School of Health Management, Binzhou Medical University, Yantai, People's Republic of China

Correspondence: Hongmei Xu, School of Nursing, Binzhou Medical University, No. 525 Huanghe 3rd Road, Binzhou, 256600, People's Republic of China, Tel +86-13754689536, Email hmx58@163.com; Yang Kong, School of Health Management, Binzhou Medical University, No. 346 Guanhai Road, Yantai, 264003, People's Republic of China, Tel +86-13773576163, Email kongyang@bzmc.edu.cn

Purpose: The purpose of this study was to establish and validate machine learning-based models for predicting the risk of venous thromboembolism (VTE) in intensive care unit (ICU) patients.

Patients and Methods: The clinical data of 1494 ICU patients who underwent Doppler ultrasonography or venography between December 2020 and March 2023 were extracted from three tertiary hospitals. The Boruta algorithm was used to screen the essential variables associated with VTE. Five machine learning algorithms were employed: Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Gradient Boosting Decision Tree (GBDT), and Logistic Regression (LR). Hyperparameter optimization was conducted on the predictive model of the training dataset. The performance in the validation dataset was measured using indicators, including the area under curve (AUC) of the receiver operating characteristic (ROC) curve, specificity, and F1 score. Finally, the optimal model was interpreted using the SHapley Additive exPlanation (SHAP) package.

Results: The incidence of VTE among the ICU patients in this study was 26.04%. We screened 19 crucial features for the risk prediction model development. Among the five models, the RF model performed best, with an AUC of 0.788 (95% CI: 0.738–0.838), an accuracy of 0.759 (95% CI: 0.709–0.809), a sensitivity of 0.633, and a Brier score of 0.166.

Conclusion: A machine learning-based model for prediction of VTE in ICU patients were successfully developed, which could assist clinical medical staff in identifying high-risk populations for VTE in the early stages so that prevention measures can be implemented to reduce the burden on the ICU patients.

Keywords: venous thromboembolism, machine learning, algorithm, prediction model, intensive care unit

Introduction

Venous thromboembolism (VTE), including deep vein thrombosis (DVT) and pulmonary embolism (PE), is a group of thromboembolic diseases and the third most prevalent cardiovascular disease after myocardial infarction and stroke.[1] Patients in the intensive care unit (ICU) are at a high risk of VTE, and the incidence of VTE is very high if sufficient preventive measures are implemented.[2,3] Several adverse outcomes are associated with VTE, including increased mortality, prolonged hospitalization, higher healthcare expenditures, decreased quality of life, and a greater risk of recurrent DVT and PE.[4,5] In the United States, deaths from VTE reach 900,000–1,000,000 per year.[6] According to a meta-analysis, patients with VTE had an average ICU stay extension of 3.92 days and an increase in mechanical ventilation of 4.85 days.[7] In a cross-sectional survey involving 907 patients with VTE, Feehan et al discovered that 40.6% of them were concerned about VTE recurrence, and 36.2% experienced severe anxiety or depression after VTE.[8] Therefore, preventing and treating VTE in the clinical setting is crucial. As a preventable illness,

3279

if patients with a high VTE risk can be identified early and accurately through risk assessment models and provided with timely intervention, it can reduce healthcare expenses and potentially enhance patient survival rates.

Recently, several models for assessing VTE risk have been developed for general inpatient, including the Padua, Caprini, and Khorana scores, which can assist clinicians in assessing the objective risk of VTE.[9–11] However, these assessment methods were not specifically designed for ICU patients and do not accurately predict VTE in critically ill patients due to the lack of specific predictive factors, including the use of particular medicines, central venous catheterization, and mechanical ventilation. Furthermore, most of the established VTE risk prediction models for ICU patients have been established using multivariate logistic regression analyses;[12–14] however, logistic regression is generally a less ideal approach when dealing with more intricate data interactions.[15] Due to the complexity and variety of severe illnesses and the volume of data generated by monitoring systems, particularly in the ICU, there is an immediate need for more data science methods and data-driven research.[16]

With the rapid development of the internet and the big data industry, machine learning algorithms are increasingly being implemented as new data analysis tools for the risk prediction of illnesses and their associated complications due to their robust computational capacity for learning. Machine learning is the science of enabling computers to learn and act like humans, improving their autonomous learning by providing them with data and information in the form of observations and real-world interactions.[17] Numerous studies have revealed the superior predictive potential of machine-learning models.[18–20] Nevertheless, these models are frequently accompanied by more intricate computation-intensive procedures with stronger predictive ability, which enhances the model's accuracy while introducing the "black box" prediction issue.[21] The future direction of artificial intelligence development must be reliable. Consequently, improving the transparency and interpretability of these models is also crucial.[22]

Thus, the purpose of this study was to establish and validate machine learning-based models that can assist in detecting and diagnosing ICU patients with a high VTE risk. Additionally, we employed an interpretable algorithm for the model to provide an intuitive elucidation of the risk associated with patient predictions, further strengthening the selected model's reliability.

Materials and Methods
Data Source
This study was designed as a retrospective study. Patient data were retrieved from the electronic medical records system (EMRS) of three tertiary hospitals: Binzhou Medical University Hospital (in Binzhou, Shandong), Yantai Affiliated Hospital of Binzhou Medical University (in Yantai, Shandong), and Binzhou People's Hospital (in Binzhou, Shandong). Critically ill patients who required ICU admission between December 2020 and March 2023 were included. In the case of multiple admissions, only the initial admission was considered. The patients who underwent color Doppler ultrasonography or venography were enrolled after applying the following inclusion criteria: i) ICU stays ≥ 48 h; ii) patients aged ≥ 18 years; and iii) possessing at least one color Doppler ultrasonography or venography screening result to identify VTE. The exclusion criteria were as follows: i) VTE (DVT or PE) occurring before or within 48 h of ICU admission; ii) more than 30% of patient data missing; and iii) diagnosis of acute leukemia. Ethics approval was obtained from the Ethics Committee of the Binzhou Medical University (No.2023383), and our study complies with the Declaration of Helsinki. Since this was only observational research and all patient data were anonymized, patient consent was waived by our institutional ethic committee. The study protocol flowchart is shown in Figure 1.

Candidate Features and Outcomes
Candidate variables were selected from the literature using the following search terms: ("venous thrombosis" or "deep vein thrombosis" or "VTE" or "DVT") and ("influence factors" or "risk factors" or "risk factor score") and ("intensive care unit" or "critically ill patients" or "critical care unit" or "ICU") across eight electronic databases: PubMed, Web of Science, Cochrane Library, Embase, CNKI, VIP, WanFang, and CBM. The retrieved literature was imported into EndNote 20 for bibliographic review. Two researchers independently screened the literature to extract risk factors and cross-checked them, guided by inclusion and exclusion criteria. Afterward, the research team deliberated on a proposal containing previous candidate variables, and candidate variables were incorporated into the final list through consensus based on their easy availability in the clinical setting and their frequency of mention in the references. Finally, 56 candidate VTE variables were determined: (1) Demographic variables:
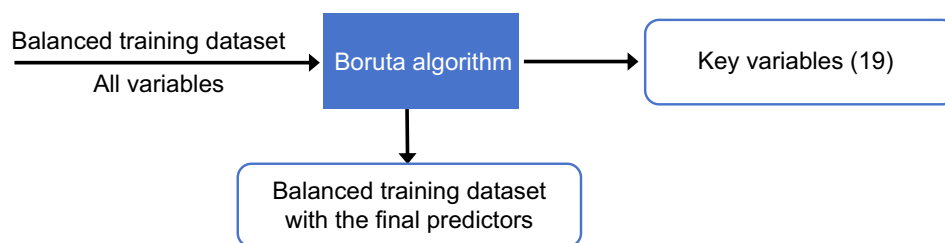
Phase 1: Data collection
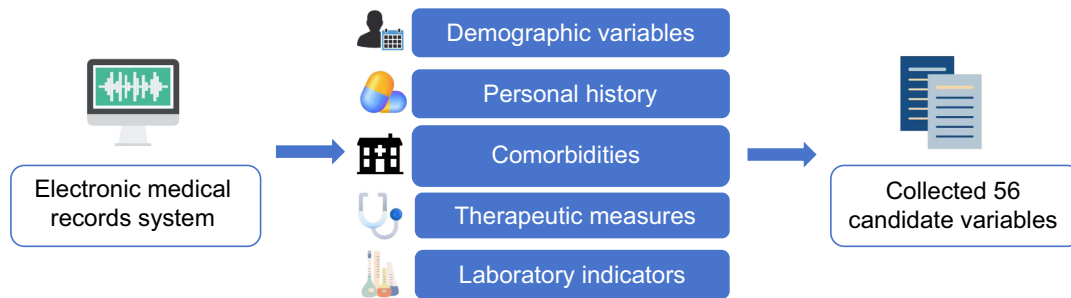
Phase 2: Data Pre-processing
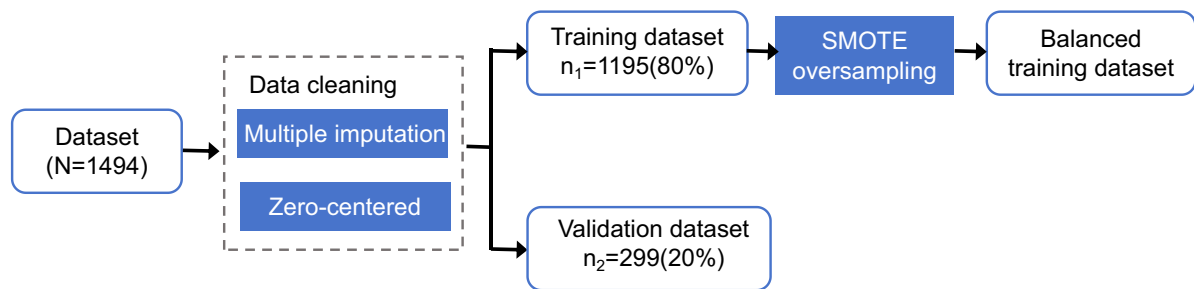
Phase 3: Feature selection

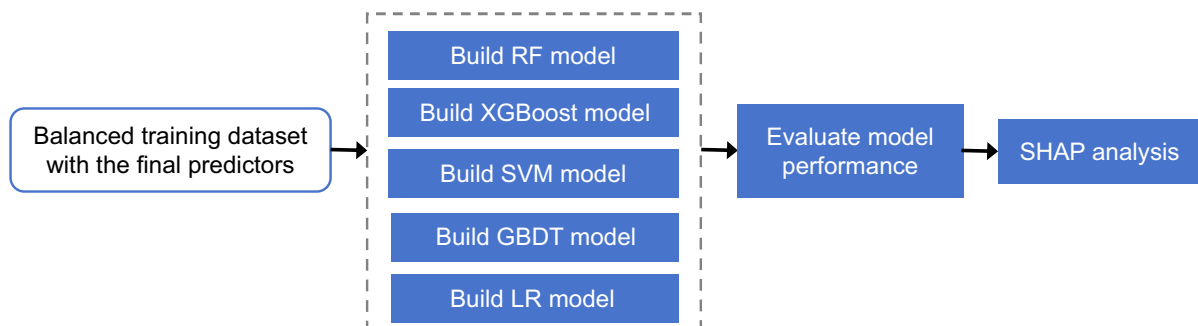**Phase 4: Model construction and validation**

Figure 1 Flowchart of the study procedure.
Abbreviations: SMOTE, Synthetic Minority Over-sampling Technique; RF, Random Forest; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; GBDT, Gradient Boosting Decision Tree; LR, Logistic Regression; SHAP, SHapley Additive exPlanation.

including age, sex, blood type, length of braking time, Glasgow Coma Scale (GCS), Acute Physiology and Chronic Health Evaluation II score (APACHE II score), disease type, muscle strength of lower limb, swelling of lower limb. (2) Personal history: including history of smoking, history of drinking, VTE history, history of coronary heart disease, history of atrial fibrillation, history of inflammatory bowel, recent surgical history, history of varicose veins in lower extremities. (3) Comorbidities: including infection, polytrauma, diabetes, hypertension, stroke, hyperlipemia, sepsis, heart failure, respiratory failure, cancer, severe acute pancreatitis, hepatic failure, rheumatic, acute myocardial infarction, end stage renal disease, serious lung disease. (4) Therapeutic measures: including sedative, vasoactive agent, transfusion of red blood cells, transfusion of platelet, central venous catheter (CVC), peripherally inserted central catheter (PICC), mechanical ventilation, mechanical ventilation time, continuous renal replacement therapy (CRRT), extracorporeal membrane oxygenation (ECMO). (5) Laboratory indicators: including c-reactive protein, red blood cell (RBC), white blood cell (WBC), platelet (PLT), hematocrit (HCT), mean platelet volume (MPV), hemoglobin (HGB), D-dimer (DD), fibrinogen (FIB), prothrombin time (PT), thrombin time (TT), activated partial thrombo-plastin time (APTT), international normalized ratio (INR).

The outcome was VTE, including DVT, PE, or both, which was objectively confirmed using color Doppler ultrasonography or venography during ICU hospitalization. These data were manually collected and double-checked from the EMRS of three tertiary hospitals using Epidata software version 3.1 by two seasoned researchers.

Data Preprocessing

Before model development, the data were preprocessed. To reduce bias due to missing data, variables with a missing rate above 30% were deleted, and missing values were imputed using multiple imputation. Multiple imputation is a highly effective and commonly employed technique for dealing with missing values, which uses correlation between variables to fill in missing data and enhances accuracy through iterative processes.[23] Continuous variables were scaled to unit variance and zero-centered to make the training process less sensitive to the scale of the variables. A computer-generated random number sequence was used to partition the entire cohort randomly into two datasets: a training dataset (80%) and a validation dataset (20%). In order to solve the problem of class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was used.[24]

Feature Selection

Feature selection is a crucial step in model construction, which can eliminate irrelevant or redundant variables to improve the model's performance and reduce computational time. This study utilizes the Boruta algorithm for feature selection within the training dataset. The Boruta algorithm is a sort wrapper algorithm based on random forest classifier that filters out all variable sets uncorrelated with the dependent variable, rather than selecting only non-redundant variables. The principle is to generate a "shadow feature" for each variable, calculate the Z-value of each variable several times using the random forest model, and then identify the importance by comparing the Z-value of each feature with the Z-value of its corresponding "shadow feature". If the Z-value of a real feature is significantly higher than the maximum Z-value of the applied shadow feature in multiple tests, the feature is labelled as important (green area), also known as an acceptable variable. Otherwise, it is labelled as "unimportant" (red area), also known as an unacceptable variable.[25]

Model Development and Validation

In this study, five machine learning algorithms, including Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Gradient Boosting Decision Tree (GBDT), and Logistic Regression (LR) were employed to construct models for predicting the risk of VTE in ICU patients. The model's hyperparameters were identified using a ten-fold cross-validation method based on grid search. Ten-fold cross-validation involves dividing the dataset into ten distinct subsets, with each subset acting as a fold. In each round of cross-validation, nine of these folds are used as the training set, while the remaining fold serves as the test set. This process is repeated ten times, ensuring that each fold serves as the test set exactly once. Cross-validation helps mitigate model overfitting and enhances its robustness. The area under curve (AUC) of the receiver operating characteristic (ROC) curve, accuracy, specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), F1-score, and Brier score were used to evaluate the models' performance in the training and validation datasets. After selecting the optimal model, we visualized the importance of the feature using the SHapley Additive exPlanation (SHAP) package.

Statistical Analysis

The samples were examined for fit to a normal distribution using the Kolmogorov–Smirnov test. Continuous variables were described using median and interquartile range (IQR), whereas categorical variables were represented by frequency and percentage. The chi-square test was used for categorical data, and the Mann–Whitney U-test was employed for continuous variables to assess the differences between the patient groups with and without VTE. A two-sided p-value < 0.05 was considered statistically significant. Statistical Package for the Social Sciences (SPSS, version 27.0), R (version 4.3.1), and Python (version 3.9) were used for statistical analysis.

Results
Baseline Characteristics

In this research involving 1494 patients, the incidence of VTE was 26.04% (389/1494), with 933 (52.40%) being male. All patients were Chinese, with a median age of 69 years, a median length of braking time of 9 days, and a median APACHE II score of 17 points. The patients were divided into VTE and non-VTE groups to determine whether they developed VTE during their ICU stay. Table 1 presents a comparison of candidate variables between the non-VTE and VTE groups.

Table 1 Demographic and Clinical Variables Between Non-VTE Group and VTE Group

Variables	Total (n=1494)	Non-VTE Group (n=1105)	VTE group (n=389)	p-value
Age (years)	69.0 (55.0, 78.0)	69.0 (53.0, 78.0)	69.0 (58.0, 78.0)	0.048
Sex (male)	933 (62.4)	677 (61.3)	256 (65.8)	0.112
Length of braking time (days)	9.0 (5.0, 16.0)	7.0 (4.0, 13.0)	17.0 (10.0, 27.0)	<0.001
GCS (points)	14.0 (6.0, 15.0)	15.0 (6.5, 15.0)	11.0 (5.0, 15.0)	<0.001
APACHE II score (points)	17.0 (12.0, 23.0)	17.0 (12.0, 23.0)	19.0 (13.0, 25.0)	0.002
Disease type				
Respiratory system	519 (34.7)	387 (35.0)	132 (33.9)	0.003
Circulatory system	215 (14.4)	166 (15.0)	49 (12.6)	
Digestive system	186 (12.4)	142 (12.9)	44 (11.3)	
Nervous system	170 (11.4)	113 (10.2)	57 (14.7)	
Urinary system	37 (2.5)	35 (3.2)	2 (0.5)	
Endocrine system	33 (2.2)	28 (2.5)	5 (1.3)	
Else	334 (22.4)	234 (21.2)	100 (25.7)	
Muscle strength of lower limb				
0	26 (1.70)	20 (1.8)	6 (1.5)	0.277
1	9 (0.60)	4 (0.4)	5 (1.3)	
2	24 (1.60)	17 (1.5)	7 (1.8)	
3	49 (3.3)	38 (3.4)	11 (2.8)	
4	38 (2.5)	24 (2.2)	14 (3.6)	
5	1348 (90.2)	1002 (90.7)	346 (88.9)	
Swelling of lower limb	222 (14.9)	165 (14.9)	57 (14.7)	0.894
History of smoking	569 (38.1)	427 (38.6)	142 (36.5)	0.455
History of drinking	267 (17.9)	198 (17.9)	69 (17.7)	0.936
VTE history	24 (1.6)	9 (0.8)	15 (3.9)	<0.001
History of coronary heart disease	309 (20.7)	231 (20.9)	78 (20.1)	0.721
History of atrial fibrillation	177 (11.8)	123 (11.1)	54 (13.9)	0.149
History of inflammatory bowel	3 (0.2)	3 (0.3)	0.0 (0.0)	0.572
Recent surgical history	244 (16.3)	148 (13.4)	96 (24.7)	<0.001
History of varicose veins in lower extremities	21 (1.4)	11 (1.0)	10 (2.6)	0.023
Infection	1022 (68.4)	737 (66.7)	285 (73.3)	0.017
Polytrauma	72 (4.8)	30 (2.7)	42 (10.8)	<0.001

(*Continued*)

Table 1 (Continued).

Variables	Total (n=1494)	Non-VTE Group (n=1105)	VTE group (n=389)	p-value
Diabetes	397 (26.6)	297 (26.9)	100 (25.7)	0.653
Hypertension	771 (51.6)	550 (49.8)	221 (56.8)	0.017
Stroke	551 (36.9)	400 (36.2)	151 (38.8)	0.357
Hyperlipemia	863 (57.8)	626 (56.7)	237 (60.9)	0.142
Sepsis	128 (8.6)	93 (8.4)	35 (9.0)	0.725
Heart failure	374 (25.0)	276 (25.0)	98 (25.2)	0.933
Respiratory failure	718 (48.1)	521 (47.1)	197 (50.6)	0.236
Cancer	252 (16.9)	192 (17.4)	60 (15.4)	0.377
Severe acute pancreatitis	28 (1.9)	20 (1.8)	8 (2.1)	0.758
Hepatic failure	105 (7.0)	77 (7.0)	28 (7.2)	0.879
Rheumatic	32 (2.1)	27 (2.4)	5 (1.3)	0.175
Acute myocardial infarction	109 (7.3)	86 (7.8)	23 (5.9)	0.223
End stage renal disease	208 (13.9)	159 (14.4)	49 (12.6)	0.380
Serious lung diseases	550 (36.8)	394 (35.7)	156 (40.1)	0.118
Sedative	1092 (73.1)	772 (69.9)	320 (82.3)	<0.001
Vaso active agent	943 (63.1)	664 (60.1)	279 (71.7)	<0.001
Transfusion of red blood cells	817 (54.7)	570 (51.6)	247 (63.5)	<0.001
Transfusion of platelet	152 (10.2)	93 (8.4)	59 (15.2)	<0.001
CVC	892 (59.7)	646 (58.5)	246 (63.2)	0.099
PICC	40 (2.7)	30 (2.7)	10 (2.6)	0.880
Mechanical ventilation	738 (49.4)	495 (44.8)	243 (62.5)	<0.001
Mechanical ventilation time (h)				
0	754 (50.5)	611 (55.3)	143 (36.8)	<0.001
0<t<96	322 (21.6)	247 (22.4)	171 (44.0)	
t≥96	418 (28.0)	247 (22.4)	75 (19.3)	
Continuous renal replacement therapy	107 (7.2)	82 (7.4)	25 (6.4)	0.513
ECMO treatment	14 (0.9)	9 (0.8)	5 (1.3)	0.601
RBC (10^{12}/L)	3.70 (3.0, 4.3)	3.7 (3.0, 4.3)	3.7 (3.0, 4.2)	0.594
WBC (10^9/L)	10.3 (7.1, 14.7)	10.2 (6.9, 14.8)	10.5 (7.3, 14.5)	0.327
PLT (10^9/L)	183.0 (120.0, 253.3)	183.0 (120.5, 255.0)	182.00 (119.0, 251.5)	0.656
HCT (%)	33.0 (27.0, 39.0)	33.0 (27.0, 39.0)	33.0 (27.0, 38.0)	0.756
MPV (fL)	10.4 (9.7, 11.4)	10.4 (9.7, 11.3)	10.5 (9.7, 11.5)	0.264
HGB (g/L)	110.5 (89.0, 130.0)	110.0 (89.5, 130.5)	111.0 (88.0, 129.0)	0.919
DD (mg/L)	2.4 (1.0, 6.0)	2.3 (1.0, 5.4)	3.3 (1.3, 8.9)	<0.001
FIB (g/L)	3.6 (2.5,4.7)	3.6 (2.6, 4.7)	3.7 (2.4, 4.7)	0.596
TT (s)	16.9 (15.7, 18.4)	16.9 (15.9, 18.4)	16.8 (15.5, 18.4)	0.083
PT (s)	13.4 (12.3, 15.0)	13.4 (12.2, 15.1)	13.5 (12.3, 14.8)	0.972
APTT (s)	33.6 (28.6, 41.5)	34.1 (29.1, 42.0)	31.8 (27.7, 39.6)	0.001
INR	1.2 (1.1, 1.3)	1.2 (1.1, 1.3)	1.2 (1.1, 1.3)	0.989

Abbreviations: GCS, Glasgow Coma Scale; APACHE II, Acute Physiology and Chronic Health Evaluation II score; VTE, venous thromboembolism; CVC, central venous catheter; PICC, peripherally inserted central catheter; ECMO, extracorporeal membrane oxygenation; RBC, red blood cell; WBC, white blood cell; PLT, platelet; HCT, hematocrit; MPV, mean platelet volume; HGB, hemoglobin; DD, D-dimer; FIB, fibrinogen; TT, thrombin time; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalized ratio.

Model Input Features

The result of feature screening based on the Boruta algorithm is shown in Figure 2. Sorted according to the Z-values, the green boxes represent variables considered important, while the red boxes indicate variables deemed unimportant. Variables shown in yellow box plots are those for which the Boruta technique could not decisively determine relevance. Ultimately, 19 essential variables were identified to build a VTE prediction model for ICU patients: length of braking time, DD, VTE history, mechanical ventilation time, APTT, recent surgical history, history of varicose veins in lower extremities, polytrauma, INR, mechanical ventilation, GCS, sedative, APACHE II score, PT, TT, HGB, RBC, HCT, and age.
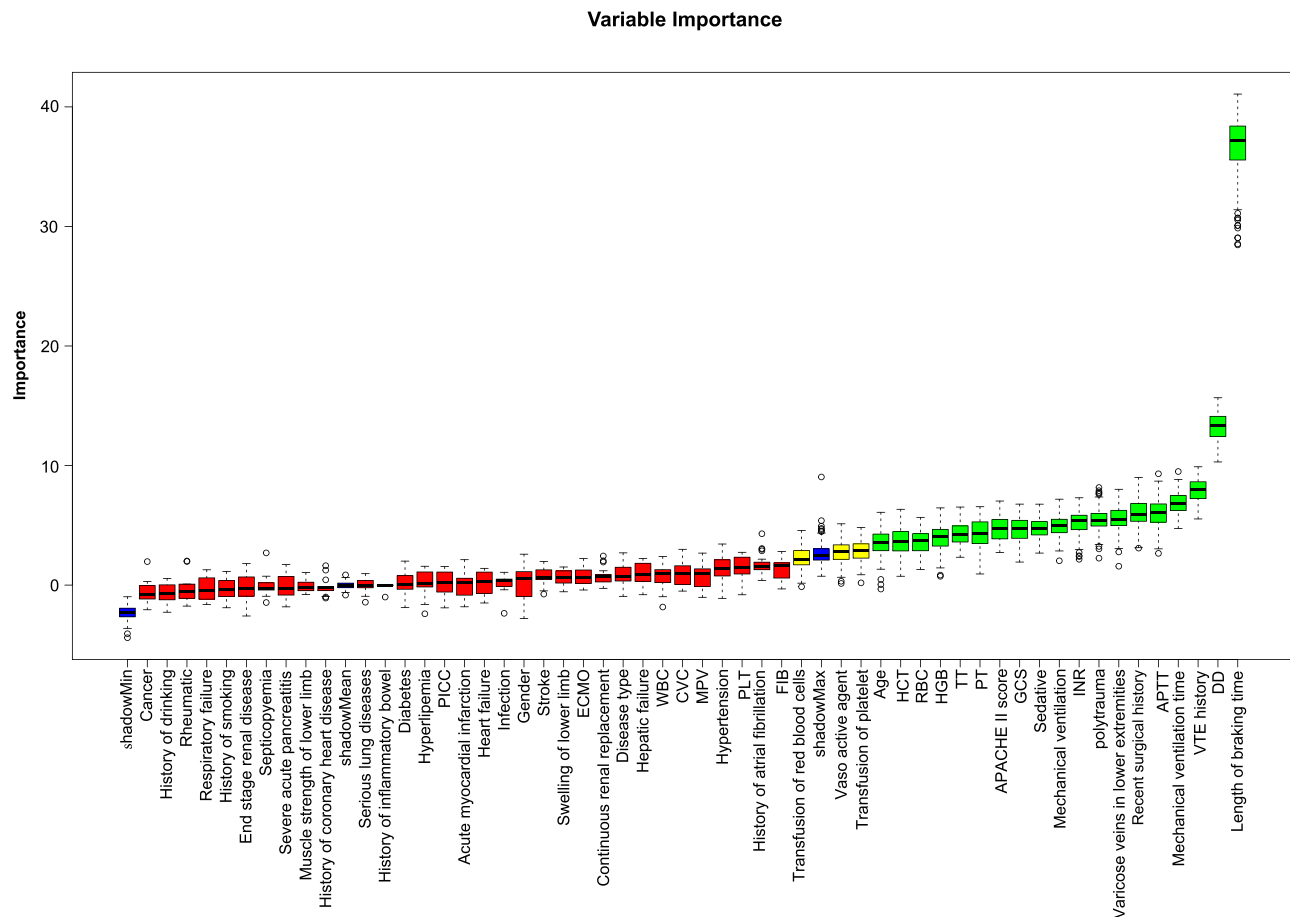
Variable Importance

Figure 2 Feature selection based on Boruta algorithm.
Abbreviations: DD, D-dimer; VTE, venous thromboembolism; APTT, activated partial thromboplastin time; INR, international normalized ratio; GCS, Glasgow Coma Scale; APACHE II score, Acute Physiology and Chronic Health Evaluation II score; PT, prothrombin time; TT, thrombin time; HGB, hemoglobin; RBC, red blood cell; HCT, hematocrit; FIB, fibrinogen; PLT, platelet; MPV, mean platelet volume; CVC, central venous catheter; WBC, white blood cell; ECMO, extracorporeal membrane oxygenation; PICC, peripherally inserted central catheter.

Model Performance

After identifying these 19 variables, five machine-learning methods were used to construct models for predicting VTE in ICU patients. To ensure optimal predictive performance for each model, we conducted thorough hyperparameter optimization in this study. The optimization of the RF parameters mainly includes the values of n_estimators (number of trees), max_depth (maximum depth per tree), max_features (maximum features considered per split), min_samples_leaf (minimum samples per leaf node), and min_samples_split (minimum samples to split an internal node). These parameters collectively influence model behavior, balancing complexity and overfitting. For SVM, we evaluated several kernels including linear, poly, sigmoid, and rbf. Ultimately, we found that the rbf kernel performed the best, as it had the lowest error. For XGBoost, optimization focused on adjusting key parameters: gamma determines the minimum loss reduction needed to split a leaf node, acting as a regularization factor; subsample controls the fraction of samples used per tree; learning_rate scales the contribution of each tree, with lower values enhancing model robustness. GBDT parameters were tuned similarly to RF. LR optimizes performance through key parameters: the penalty parameter selects the type of regularization to apply; the C parameter controls the strength of regularization, with smaller values promoting better generalization to new data; and the solver selects the optimization method for the objective function. The best-tuned hyperparameters for all models are listed in Table 2. In the training set, the predictive value of the model was assessed using ten-fold cross-validation, as shown in Figure 3. RF exhibited the highest clinical predictive value, achieving an AUC of 1.000 (95% CI: 0.9777–1.023), followed by GBDT, with an AUC of 0.998 (95% CI 0.975–1.021).

Table 2 The Best-Tuned Hyperparameters for Each Model

Models	Hyperparameters
RF	max_depth=15, max_features=4, min_samples_split=2, min_samples_leaf=1, n_estimators=67
XGBoost	max_depth=4, n_estimators=94, gamma=0, subsample=1, learning_rate=0.2
SVM	C=1.4, gamma=0.4, kernel='rbf', probability=True
GBDT	max_depth=4, max_features='sqrt', min_samples_split=2, n_estimators=90, learning_rate=0.2, subsample=1
LR	C=0.001, penalty='l2', solver=liblinear

Abbreviations: RF, Random Forest; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; GBDT, Gradient Boosting Decision Tree; LR, Logistic Regression.

Furthermore, we verified the five predictive models in the validation dataset for stability and generalizability. Table 3 provides an overview of each model's performance. The ROC curves are displayed in Figure 4A, with AUCs varying from 0.709 (95% CI: 0.659–0.759) to 0.789 (95% CI: 0.738–0.838). The calibration plots of the five models are presented in Figure 4B. Among the five models, the RF model outperformed the other models in terms of accuracy, PPV, and F1 value. The Brier score of the RF model for predicting VTE was 0.166, demonstrating the model's reliability. The GBDT model had a slightly higher AUC than RF, but its lower sensitivity indicates difficulty in correctly identifying positive cases. Although the LR model had higher sensitivity than the RF model, its low specificity resulted in too many non-VTE patients being incorrectly predicted as VTE patients. Additionally, its relatively high Brier score of 0.217 suggests that the LR model's probabilistic predictions might be biased compared to actual outcomes. Therefore, the RF model was ultimately selected as the optimal model for this study.

Explanation of the RF Model

To visually illustrate the importance of selected variables, we utilized SHAP package to analyze their contributions to the output results of the RF model. Figure 5A depicts the ranking of importance among the 19 variables in the predictive model, ordered from top to bottom based on decreasing levels of importance. The top five predictor variables that contributed most to the prediction of outcome were the length of braking time, DD, APTT, mechanical ventilation time and age. Figure 5B illustrates the positive and negative effects of features on the predicted value of the model. Each dot in the graph represents a sample, with each row representing a variable. Points with higher red saturation indicate larger values for that variable in relation to a specific patient, while points with higher blue saturation denote smaller values. The horizontal axis represents the SHAP value itself: the greater its absolute value, the more influential the variable is. Dots positioned towards the right side indicate that higher values of the feature positively contribute more to the
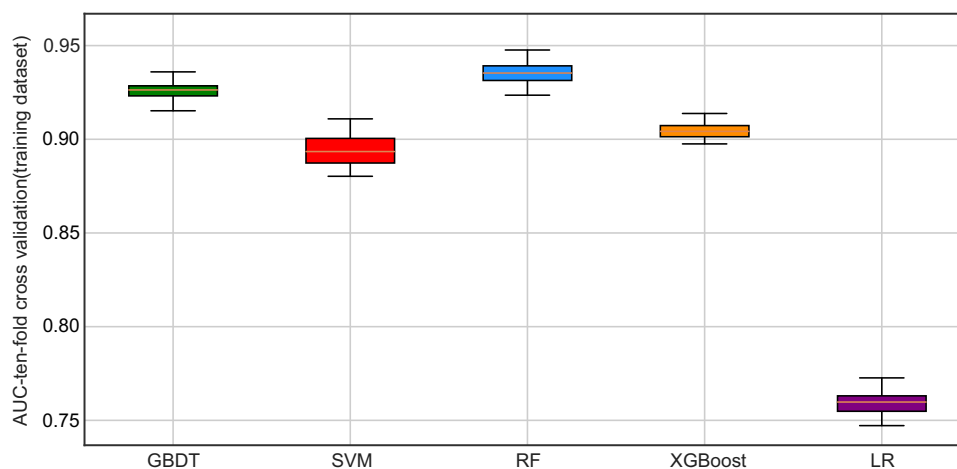
Figure 3 Clinical predictive value of five machine learning models (ten-fold cross-validation) in the training dataset.
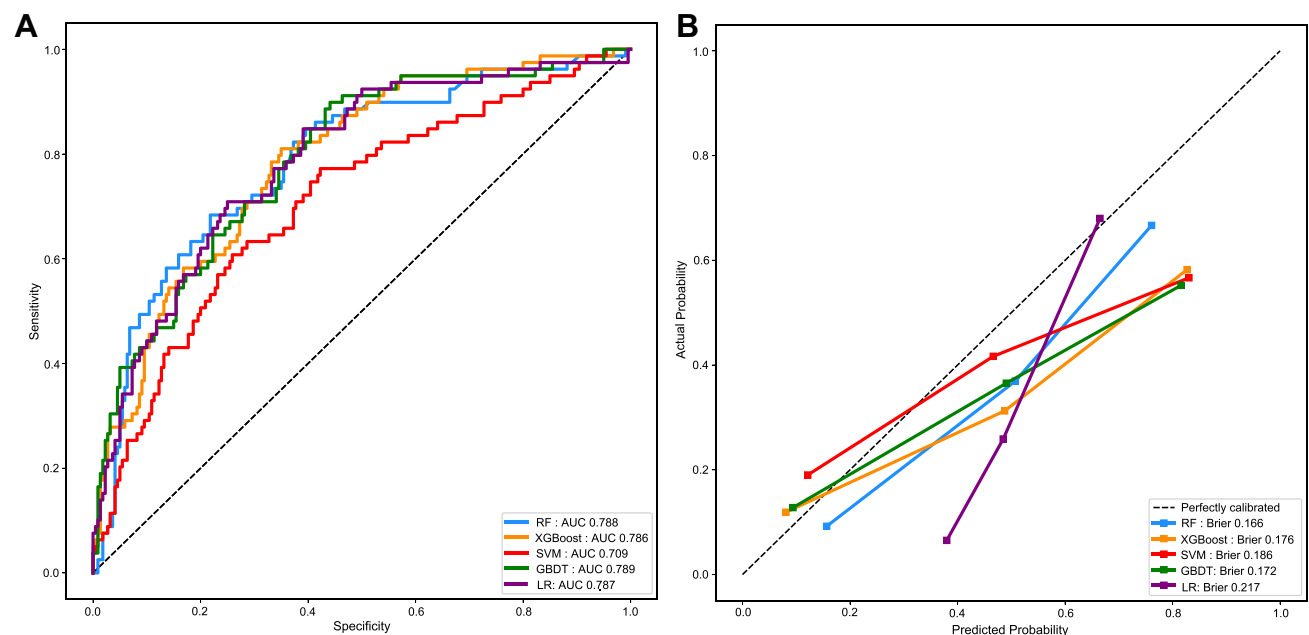**Abbreviations**: AUC, area under the curve; GBDT, Gradient Boosting Decision Tree; SVM, Support Vector Machine; RF, Random Forest; XGBoost, eXtreme Gradient Boosting; LR, Logistic Regression.

Table 3 Model Performance in Predicting VTE in the Training and Validation Datasets

	AUC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity	PPV	NPV	F1	Brier
Training dataset								
RF	1.000 (0.977, 1.023)	0.999 (0.976, 1.023)	1.000	0.999	0.999	1.000	0.999	0.021
XGBoost	0.997 (0.973, 1.020)	0.974 (0.951, 0.997)	0.985	0.963	0.964	0.985	0.974	0.037
SVM	0.994 (0.971, 1.018)	0.971 (0.948, 0.994)	0.975	0.967	0.967	0.975	0.971	0.023
GBDT	0.998 (0.975, 1.021)	0.979 (0.956, 1.002)	0.985	0.973	0.973	0.985	0.979	0.030
LR	0.768 (0.745, 0.791)	0.680 (0.656, 0.703)	0.599	0.760	0.714	0.655	0.652	0.219
Validation dataset								
RF	0.788 (0.738, 0.838)	0.759 (0.709, 0.809)	0.633	0.805	0.538	0.859	0.581	0.166
XGBoost	0.786 (0.736, 0.837)	0.722 (0.672, 0.772)	0.608	0.764	0.480	0.844	0.536	0.176
SVM	0.709 (0.659, 0.759)	0.742 (0.692, 0.793)	0.367	0.877	0.518	0.794	0.430	0.186
GBDT	0.789 (0.739, 0.840)	0.736 (0.686, 0.786)	0.582	0.791	0.500	0.841	0.538	0.172
LR	0.787 (0.737, 0.837)	0.726 (0.676, 0.776)	0.709	0.732	0.487	0.875	0.577	0.217

Abbreviations: AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; RF, Random Forest; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; GBDT, Gradient Boosting Decision Tree; LR, Logistic Regression.

assessment of VTE, whereas dots towards the left side suggest that lower values contribute more negatively to this assessment. Regarding age, most red (aged patients) sample points were located on the right half of the vertical axis, and the corresponding SHAP value was positive, positively affecting the model's prediction of the occurrence of VTE. Therefore, it may be demonstrated that among ICU patients, older age may correspond to a higher risk of VTE. In addition, we randomly selected two patients for personalized analysis based on the model's predictions of VTE and non-VTE cases, respectively. In Figure 5C and D, the red line indicates that the feature positively contributes to the prediction, while the blue line indicates negative contribution, with longer lines indicating greater influence. The base value, represented by the average of predictions across all samples, serves as the baseline prediction of VTE for each sample. The value outputted by f(x) represents the predicted value of each sample when individual features are taken into account. For each individual patient, the model predicts the occurrence or non-occurrence of VTE based on different

Figure 4 (**A**) ROC curves of five machine learning models in validation dataset. (**B**) Calibration curves of five machine learning models in validation dataset.
Abbreviations: RF, Random Forest; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; GBDT, Gradient Boosting Decision Tree; LR, Logistic Regression.
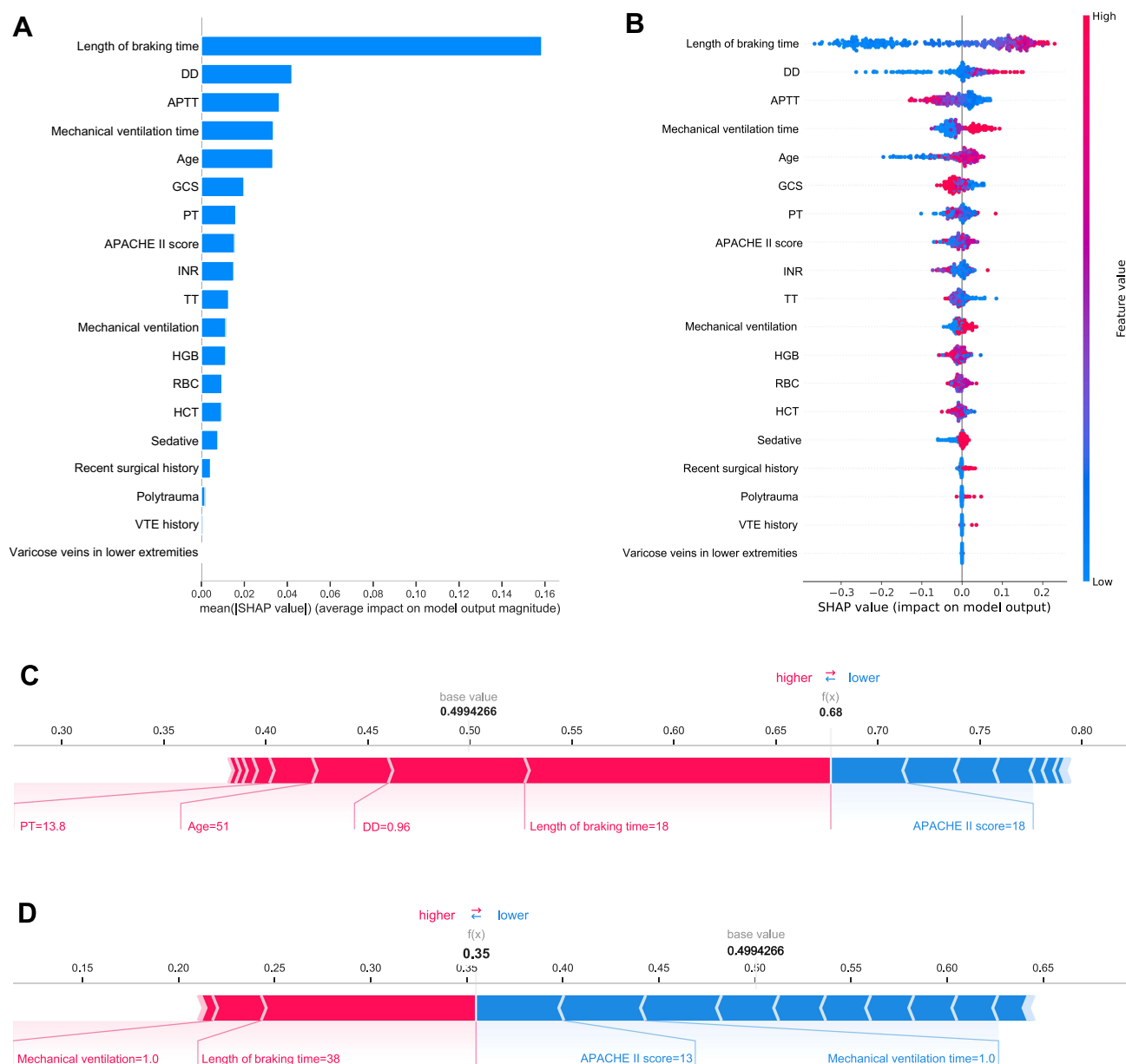
Figure 5 (**A**) Feature importance ranking as indicated by SHAP. (**B**) Attributes of features in SHAP. (**C**) The SHAP value force plot of a VTE patient was used to individually predict the characteristic variables. (**D**) The value force plot of a non-VTE patient was used to individually predict the characteristic variables.
Abbreviations: DD, D-dimer; APTT, activated partial thromboplastin time; GCS, Glasgow Coma Scale; APACHE II score, Acute Physiology and Chronic Health Evaluation II score; INR, international normalized ratio; TT, thrombin time; HGB, hemoglobin; RBC, red blood cell; HCT, hematocrit; VTE, venous thromboembolism.

reasons, and the final result is a combination of several variables that each patient possesses as a unique individual. As depicted in Figure 5C, the predicted base value for this sample is 0.5. When considering the combined effect of each feature, the predicted value for this sample is 0.68. Consequently, the model predicts that this patient is at risk of VTE, which aligns with the actual occurrence of VTE in this case.

Discussion

In addition to the common risk factors observed in the general inpatient population, ICU patients are susceptible to ICU-specific risk factors, such as invasive manipulations (central venous catheter implantation), mechanical ventilation, and the use of sedative.[26–28] These factors significantly elevate the risk of VTE in ICU patients, a condition associated with

higher patient mortality rates and increased economic burden. Early identification and management are crucial in preventing VTE in ICU patients. Currently, there are no specific VTE risk assessment tools validated specifically for ICU patients. Clinicians often rely on more generalized risk assessment tools like the Caprini and Padua scores. However, there are limitations in the sensitivity and specificity of these tools. Based on the study design and applicability to the study population, the clinical generalizability of each assessment model to ICU patients is yet to be demonstrated. Original studies frequently employed univariate filtering to screen for significant factors, followed by multiple logistic regression to identify independent risk factors for VTE to construct simple prediction models.[29,30] The development of the internet and the big data industry, particularly in the ICU, has reduced the accuracy of conventional statistical methods due to the intricate nature of patient conditions. This study proposes the use of machine learning algorithms to predict VTE risk in ICU patients, providing personalized predictions to guide clinical grading management.

Our study comprised four main steps: data preprocessing, feature selection, model construction and validation, and visualization and analysis of the model. We utilized a grid search method for automated hyperparameter optimization and performed ten-fold cross-validation to determine the best-performing configuration for each algorithm. After evaluating the performance of all constructed prediction models, we concluded that the RF model exhibited the highest performance, with 0.788 AUC, 0.805 specificity, 0.633 sensitivity, and 0.166 Brier score.

RF, GBDT, and XGBoost are all ensemble learning algorithms based on decision trees. They excel in handling classification and regression tasks, demonstrating strong predictive performance in various studies, but there are some significant differences in the way they work. RF achieves ensemble integration by training multiple decision trees independently. Each tree is built on a randomly selected subset of the dataset, and at each node of the tree, a random subset of features is considered for splitting.[31] GBDT integrates by sequentially training decision trees, where each subsequent tree is constructed based on the predictions of the ensemble from previous trees. The main goal is to minimize the gradient of a specified loss function. This iterative process continuously enhances the predictive power of the model.[32] XGBoost represents a significant enhancement and optimization of GBDT. Compared to GBDT, XGBoost introduces several key improvements. The most notable change is in its approach to the loss function: while GBDT primarily minimizes the negative gradient of the loss function, XGBoost leverages the second-order Taylor expansion. Additionally, XGBoost enhances model robustness through the incorporation of a regularization term in its objective function.[33] This regularization helps manage the complexity of the model by penalizing overly complex trees, thereby mitigating overfitting and improving generalization performance. SVM make use of the geometric relationships of variables to separate different classes of data points for outcome prediction by finding an optimal hyperplane in the feature space. It performs well in binary classification problems and provides flexibility for both nonlinear and linear problems.[34] However, as the dataset grows larger, SVM experience a notable increase in training time. Moreover, SVM has poor interpretability.[35] LR typically performs well in probabilistic prediction tasks. It predetermines the association between the predictor variables in a linear fashion, which gives it the ability to explain each variable causally.[36] However, if there is a non-linear relationship between the variables in the dataset, LR may not be the optimal choice.[37]

While the RF model demonstrates superior performance, its inherent black-box nature could restrict its application. To increase the interpretability and transparency of the model, we conducted a SHAP analysis on the best-performing RF model, which helps improve patients' and physicians' understanding of the decision-making process. Previous studies have established that age and some comorbidities are significant factors for VTE in ICU patients. Our results align with those of earlier research. Patients with prolonged braking time possess a higher chance of VTE because of an increased probability of encountering medical complications and functional impairments.[38] Aging is correlated with an increased risk of VTE. Potential causes of venous stasis include the cumulative impact of risk factors on the venous wall, less frequent activity, increased immobility, and high blood thrombin levels.[39,40] Similar to prior findings, our results demonstrated that recent surgery or polytrauma may increase the risk of VTE.[41,42] Body trauma may lead to the development of hypoxemia, and hypoxia after oxygenation may induce tissue damage.[43] Additionally, our findings indicated that patients were more susceptible to VTE if they had higher APACHE II scores and lower GCS scores. The GCS score is frequently used to determine coma level; higher coma severity corresponds with lower score. The higher the APACHE II score, the greater the disease severity and decrease in organ function. Furthermore, individuals with VTE

history stood a strong chance of getting VTE once more, indicating that the possibility of VTE events in these patients remains existent.

Therapeutic interventions, including sedative and mechanical ventilation, are linked to an increased risk of VTE. Mechanical ventilation is a common form of supportive therapy for ICU patients that involves replacing or controlling a patient's respiratory movements with a mechanical device. Prior research have confirmed that mechanical ventilation is an independent risk factor for VTE development.[44] Mechanical ventilation and positive end-expiratory pressure increase the right ventricular load, decrease the left ventricular load and total output, and increase the incidence of venous hemosiderosis, while mechanical ventilation also alters pulmonary fibrin conversion and increases the incidence of coagulation.[45] Simultaneously, our study showed that the risk of VTE increases with length of ventilation. Sedative leads to a decrease in venous blood flow velocity, which aggravates blood stasis and increases VTE incidence.

Our findings also indicated higher D-dimer and lower APTT, TT, and PT, similarly increasing the VTE risk. D-dimer, a specific hydrolysis product of cross-linked fibrin generated by fibrinolytic enzymes, can serve as a marker for secondary hyperfibrinolysis and the hypercoagulable condition of blood.[46,47] Many investigations have established a correlation between elevated D-dimer and an increased chance of thrombosis. APTT is a coagulation test commonly used to evaluate the blood coagulation system's functional status. A shortened APTT signifies that the blood is in a hypercoagulability state, which predisposes VTE development.[48] Lower PT levels are associated with an increased risk of VTE, possibly as a result of coagulation factors II, V, VII, and X becoming more active.[49] Along with its derived measures of INR, these assays evaluate the extrinsic pathway of coagulation.

In this study, we constructed prediction models using multiple machine-learning algorithms. By examining the AUC, sensitivity, specificity, and calibration curves, we discovered that the RF model outperformed the other models. The optimal model can facilitate clinicians to identify high-risk groups of VTE early, implement preventative therapies in time to lower the incidence of VTE and improve patient prognosis. Simultaneously, this approach helped alleviate the economic strain on patients, particularly those with less favorable financial circumstances. The most effective use of the model was to integrate it into an electronic health record system. The patient's medical data are gradually uploaded to the system without human resources, enabling automated, real-time monitoring to facilitate intelligent early warnings of VTE in clinical practice. Our machine learning models were derived from a dataset of 1494 individuals from three tertiary hospitals, giving our model some reliability and generalizability. Furthermore, we applied the SHAP method to the RF model to achieve better interpretability and to assist healthcare professionals and patients in better understanding the model's decision-making process.

However, there are some limitations. First, clinical data collection is not sufficiently comprehensive, and potential predictors may be overlooked. Second, this study just utilized data from the patient's initial admission to the ICU and failed to reflect the time series characteristics of the ICU, which may provide more accurate information. Third, our study lacks an independent dataset for external validation; therefore, the model's performance when applied to an additional dataset is unknown. In the future, we plan to collect sufficient data to test the model's performance in real-world scenarios. To further optimize the model's accuracy, we will pursue its ongoing testing and improvement in clinical practice.

Conclusion

In summary, we constructed and validated several predictive models for VTE in ICU patients using machine-learning algorithms, and the RF model showed better performance. This effective computer-aided approach can potentially help predict the occurrence of VTE in ICU and enhance the patient's survival rate. Furthermore, this research explains the significance of key risk factors in forecasting outcomes through SHAP analysis, thereby helping physicians identify predictors and improve the reliability of the projected results.

Data Sharing Statement

The data used to support the findings of this study are available from the corresponding author (Hongmei Xu) upon request.

Acknowledgments

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Previtali E, Bucciarelli P, Passamonti SM, et al. Risk factors for venous and arterial thrombosis. *Blood Transfus*. 2011;9(2):120–138. doi:10.2450/2010.0066-10
2. Cook D, Douketis J, Crowther MA, et al. The diagnosis of deep venous thrombosis and pulmonary embolism in medical-surgical intensive care unit patients. *J Crit Care*. 2005;20(4):314–319. doi:10.1016/j.jcrc.2005.09.003
3. Cook D, Meade M; PROTECT Investigators for the Canadian Critical Care Trials Group and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Dalteparin versus unfractionated heparin in critically ill patients. *N Engl J Med*. 2011;364(14):1305–1314. doi:10.1056/NEJMoa1014475
4. Heit JA. Epidemiology of venous thromboembolism. *Nat Rev Cardiol*. 2015;12(8):464–474. doi:10.1038/nrcardio.2015.83
5. Henke PK, Kahn SR, Pannucci CJ, et al. Call to action to prevent venous thromboembolism in hospitalized patients: a policy statement from the American Heart Association. *Circulation*. 2020;141(24):e914–e931. doi:10.1161/CIR.0000000000000769
6. Centers for Disease Control and Prevention. Impact of Blood Clots on the United States; 2018.
7. Malato A, Dentali F, Siragusa S, et al. The impact of deep vein thrombosis in critically ill patients: a meta-analysis of major clinical outcomes. *Blood Transfus*. 2015;13(4):559. doi:10.2450/2015.0277-14
8. Feehan M, Walsh M, Van Duker H, et al. Prevalence and correlates of bleeding and emotional harms in a national US sample of patients with venous thromboembolism: a cross-sectional structural equation model. *Thromb Res*. 2018;172:181–187. doi:10.1016/j.thromres.2018.05.025
9. Zhou H, Hu Y, Li X, et al. Assessment of the risk of venous thromboembolism in medical inpatients using the Padua prediction score and Caprini risk assessment model. *J Atheroscler Thromb*. 2018;25(11):1091–1104. doi:10.5551/jat.43653
10. Bartlett MA, Mauck KF, Stephenson CR, et al. Perioperative venous thromboembolism prophylaxis. *Mayo Clin Proc*. 2020;95(12):2775–2798. doi:10.1016/j.mayocp.2020.06.015
11. Mulder FI, Candeloro M, Kamphuisen PW, et al. The Khorana score for prediction of venous thromboembolism in cancer patients: a systematic review and meta-analysis. *Haematologica*. 2019;104(6):1277–1287. doi:10.3324/haematol.2018.209114
12. Lin J, Zhang Y, Lin W, et al. Development and Validation of a Risk Assessment Model for Venous Thromboembolism in Patients with Invasive Mechanical Ventilation. *Cureus*. 2022;14(7):e27164. doi:10.7759/cureus.27164
13. Meizoso JP, Karcutskie CA, Ray JJ, et al. A simplified stratification system for venous thromboembolism risk in severely injured trauma patients. *J Surg Res*. 2017;207:138–144. doi:10.1016/j.jss.2016.08.072
14. Qi Z, Ding D, Wu C, et al. Construction and validation of a predictive model for early occurrence of lower extremity deep venous thrombosis in ICU patients with sepsis. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*. 2024;36(5):471–477. doi:10.3760/cma.j.cn121430-20231117-00985
15. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593
16. Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest*. 2018;154(5):1239–1248. doi:10.1016/j.chest.2018.04.037
17. Mishra A, Ashraf MZ. Using artificial intelligence to manage thrombosis research, diagnosis, and clinical management. *Semin Thromb Hemost*. 2020;46(4):410–418. doi:10.1055/s-0039-1697949
18. Ferroni P, Zanzotto FM, Scarpato N, et al. Risk assessment for venous thromboembolism in chemotherapy-treated ambulatory cancer patients. *Med Decis Making*. 2017;37(2):234–242. doi:10.1177/0272989X16662654
19. Sabra S, Mahmood Malik K, Alobaidi M. Prediction of venous thromboembolism using semantic and sentiment analyses of clinical narratives. *Comput Biol Med*. 2018;94:1–10. doi:10.1016/j.compbiomed.2017.12.026
20. Wang X, Yang YQ, Liu SH, et al. Comparing different venous thromboembolism risk assessment machine learning models in Chinese patients. *J Eval Clin Pract*. 2020;26(1):26–34. doi:10.1111/jep.13324
21. Molnar C, Casalicchio G, Bischl B Interpretable machine learning–a brief history, state-of-The-art and challenges. InJoint European Conference on Machine Learning and Knowledge Discovery in Databases; 2020:417–431.
22. Smuha NA. The EU approach to ethics guidelines for trustworthy artificial intelligence. *Comp Law Rev Int*. 2019;20(4):97–106. doi:10.9785/cri-2019-200402

23. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med.* 2016;4(2):30. doi:10.3978/j.issn.2305-5839.2015.12.63

24. Sun R, Wang X, Jiang H, et al. Prediction of 30-day mortality in heart failure patients with hypoxic hepatitis: development and external validation of an interpretable machine learning model. *Front Cardiovasc Med.* 2022;9:1035675. doi:10.3389/fcvm.2022.1035675

25. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw.* 2010;36(11):1–13. doi:10.18637/jss.v036.i11

26. Baskin JL, Pui CH, Reiss U, et al. Management of occlusion and thrombosis associated with long-term indwelling central venous catheters. *Lancet.* 2009;374(9684):159–169. doi:10.1016/S0140-6736(09)60220-8

27. Minet C, Potton L, Bonadona A, et al. Venous thromboembolism in the ICU: main characteristics, diagnosis and thromboprophylaxis. *Crit Care.* 2015;19(1):287. doi:10.1186/s13054-015-1003-9

28. Cook D, Attia J, Weaver B, et al. Venous thromboembolic disease: an observational study in medical-surgical intensive care unit patients. *J Crit Care.* 2000;15(4):127–132. doi:10.1053/jcrc.2000.19224

29. Viarasilpa T, Panyavachiraporn N, Marashi SM, et al. Prediction of symptomatic venous thromboembolism in critically ill patients: the ICU-venous thromboembolism score. *Crit Care Med.* 2020;48(6):e470–e479. doi:10.1097/CCM.0000000000004306

30. McCurdy JD, Israel A, Hasan M, et al. A clinical predictive model for post-hospitalisation venous thromboembolism in patients with inflammatory bowel disease. *Aliment Pharmacol Ther.* 2019;49(12):1493–1501. doi:10.1111/apt.15286

31. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32. doi:10.1023/A:1010933404324

32. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat.* 2001;29(5):1189–1232. doi:10.1214/aos/1013203451

33. Chen T, Guestrin C XGBoost: a scalable tree boosting system. InProceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining-KDD. San Francisco, CA, USA; 2016.

34. Deka PC, Deka PC. Support vector machine applications in the field of hydrology: a review. *Appl Soft Comput.* 2014;19:372–386. doi:10.1016/j.asoc.2014.02.002

35. Zhang XH, Heller KA, Hefter I, et al. Sequence information for the splicing of human pre-mRNA identified by support vector machine classification. *Genome Res.* 2003;13(12):2637–2650. doi:10.1101/gr.1679003

36. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform.* 2002;35(5–6):352–359. doi:10.1016/s1532-0464(03)00034-0

37. Work JW, Ferguson JG, Diamond GA. Limitations of a conventional logistic regression model based on left ventricular ejection fraction in predicting coronary events after myocardial infarction. *Am J Cardiol.* 1989;64(12):702–707. doi:10.1016/0002-9149(89)90751-0

38. Lung BE, Kanjiya S, Bisogno M, et al. Risk factors for venous thromboembolism in total shoulder arthroplasty. *JSES Open Access.* 2019;3 (3):183–188. doi:10.1016/j.jses

39. Hirmerova J, Seidlerova J, Subrt I. Deep vein thrombosis and/or pulmonary embolism concurrent with superficial vein thrombosis of the legs: cross-sectional single center study of prevalence and risk factors. *Int Angiol.* 2013;32(4):410–416.

40. Rumley A, Emberson JR, Wannamethee SG, et al. Effects of older age on fibrin D-dimer, C-reactive protein, and other hemostatic and inflammatory variables in men aged 60–79 years. *J Thromb Haemost.* 2006;4(5):982–987. doi:10.1111/j.1538-7836.2006.01889.x

41. Mi YH, Xu MY. Trauma-induced pulmonary thromboembolism: what's update? *Chin J Traumatol.* 2022;25(2):67–76. doi:10.1016/j.cjtee.2021.08.003

42. Nemeth B, Lijfering WM, Nelissen RGHH, et al. Risk and Risk Factors Associated with Recurrent Venous Thromboembolism Following Surgery in Patients With History of Venous Thromboembolism. *JAMA Network Open.* 2019;2(5):e193690. doi:10.1001/jamanetworkopen.2019.3690

43. Owens AP, Mackman N. Tissue factor and thrombosis: the clot starts here. *Thromb Haemost.* 2010;104(03):432–439. doi:10.1160/TH09-11-0771

44. Britos M, Smoot E, Liu KD, et al. The value of positive end-expiratory pressure and Fio₂ criteria in the definition of the acute respiratory distress syndrome. *Crit Care Med.* 2011;39(9):2025–2030. doi:10.1097/CCM.0b013e31821cb774

45. Cook D, Crowther M, Meade M, et al. Deep venous thrombosis in medical-surgical critically ill patients: prevalence, incidence, and risk factors. *Crit Care Med.* 2005;33(7):1565–1571. doi:10.1097/01.ccm.0000171207.95319.b2

46. Weitz JI, Fredenburgh JC, Eikelboom JW. A Test in Context: d-Dimer. *J Am Coll Cardiol.* 2017;70(19):2411–2420. doi:10.1016/j.jacc.2017.09.024

47. Favresse J, Lippi G, Roy PM, et al. D-dimer: preanalytical, analytical, postanalytical variables, and clinical applications. *Crit Rev Clin Lab Sci.* 2018;55(8):548–577. doi:10.1080/10408363.2018.1529734

48. Tripodi A, Chantarangkul V, Martinelli I, et al. A shortened activated partial thromboplastin time is associated with the risk of venous thromboembolism. *Blood.* 2004;104(12):3631–3634. doi:10.1182/blood-2004-03-1042

49. Dorgalaleh A, Daneshi M, Rashidpanah J, et al. An overview of hemostasis. In: *Congenital Bleeding Disorders: Diagnosis and Management.* Springer; 2018:3–26. doi:10.1007/978-3-319-76723-9_1