

RESEARCH

Open Access



Identification of key risk factors for venous thromboembolism in urological inpatients based on the Caprini scale and interpretable machine learning methods

Chao Liu^{1,2†}, Wei-Ying Yang^{3†}, Fengmin Cheng³, Ching-Wen Chien^{2*}, Yen-Ching Chuang^{4,5,6*} and Yanjun Jin^{7*}

Abstract

Purpose To identify the key risk factors for venous thromboembolism (VTE) in urological inpatients based on the Caprini scale using an interpretable machine learning method.

Methods VTE risk data of urological inpatients were obtained based on the Caprini scale in the case hospital. Based on the data, the Boruta method was used to further select the key variables from the 37 variables in the Caprini scale. Furthermore, decision rules corresponding to each risk level were generated using the rough set (RS) method. Finally, random forest (RF), support vector machine (SVM), and backpropagation artificial neural network (BPANN) were used to verify the data accuracy and were compared with the RS method.

Results Following the screening, the key risk factors for VTE in urology were "(C₁) Age," "(C₂) Minor Surgery planned," "(C₃) Obesity (BMI > 25)," "(C₈) Varicose veins," "(C₉) Sepsis (< 1 month)," "(C₁₀) Serious lung disease incl. pneumonia (< 1 month)" "(C₁₁) COPD," "(C₁₆) Other risk," "(C₁₈) Major surgery (> 45 min)," "(C₁₉) Laparoscopic surgery (> 45 min)," "(C₂₀) Patient confined to bed (> 72 h)," "(C₁₈) Malignancy (present or previous)," "(C₂₃) Central venous access," "(C₃₁) History of DVT/PE," "(C₃₂) Other congenital or acquired thrombophilia," and "(C₃₄) Stroke (< 1 month." According to the decision rules of different risk levels obtained using the RS method, "(C₁) Age," "(C₁₈) Major surgery (> 45 minutes)," and "(C₂₁) Malignancy (present or previous)" were the main factors influencing mid- and high-risk levels, and some suggestions on VTE prevention were indicated based on these three factors. The average accuracies of the RS, RF, SVM, and BPANN models were 79.5%, 87.9%, 92.6%, and 97.2%, respectively. In addition, BPANN had the highest accuracy, recall, F1-score, and precision.

[†]Chao Liu and Wei-Ying Yang contributed equally to this study.

*Correspondence:

Ching-Wen Chien

ihhca@sz.tsinghua.edu.cn

Yen-Ching Chuang

yenching.chuang@gmail.com; ycchuang@tzc.edu.cn

Yanjun Jin

jinyj@enzemed.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusions The RS model achieved poorer accuracy than the other three common machine learning models. However, the RS model provides strong interpretability and allows for the identification of high-risk factors and decision rules influencing high-risk assessments of VTE in urology. This transparency is very important for clinicians in the risk assessment process.

Keywords Venous thromboembolism, Caprini, Risk, Machine learning, Random forest, Boruta, Rough set

Introduction

Venous thromboembolism (VTE) is a common surgical complication in urology [1]. It is a type of occlusive venous disease caused by abnormal coagulation of venous blood, such as deep venous thrombosis (DVT) and pulmonary thromboembolism [2]. VTE can lead to death or related health damage, prolonged hospitalization, and increased treatment costs and has become a major acquired disease in hospitals [2, 3]. According to a multicenter study in China, the mortality rate of patients hospitalized for VTE doubled between 2007 and 2016 [3]. With the rapid development of minimally invasive techniques, most urological diseases are treated with minimally invasive surgery, such as laparoscopy, nephroscopy, ureteroscopy, and cystoscopy [4]. Compared with traditional laparotomy, minimally invasive surgery significantly reduces the incidence of VTE. Previous studies have reported that the incidence of DVT in patients who undergo urological laparoscopy is 0.7–10.3% [5]. Therefore, the degree of harm caused by VTE has received increasing attention from clinicians, hospital administrators, and health economics researchers.

According to the American Society of Hematology in 2018 [6], the VTE Management Guidelines issued by the European Respiratory Society in 2019 [7], and an epidemiological study of VTE in the Chinese population [8], effective preventive measures can significantly reduce the incidence of VTE. For example, Fernando, Tran [9] found that using appropriate drugs and mechanical prevention could reduce the incidence of DVT. Hussain, Kim [10] reported that within 180 days, 61% of hospitalized patients who were primarily diagnosed with VTE had not received effective preventive measures against VTE at first admission, leading to subsequent VTE. Therefore, VTE risk should be assessed to accurately predict the degree of VTE risk, formulate appropriate prevention strategies, reduce the incidence of VTE and VTE-related mortality, and improve the surgical quality of urological patients.

The Caprini scale is a VTE risk assessment system developed by Caprini based on clinical data and experience [11]. Notably, some studies have shown that the Caprini scale is an effective and feasible tool for assessing VTE risk in urological inpatients [12, 13]. For example, K, T [12] reported that in all majors, the risk of VTE increased significantly in patients with a Caprini score ≥ 5 , and more surgical majors used the Caprini

score to predict VTE. Frankel, Belanger [13] reported that the Caprini score was a significant independent predictor of VTE in patients undergoing robot-assisted radical prostatectomy.

The effectiveness of the Caprini scale for assessing VTE risk has been extensively studied and confirmed. However, for hospitalized patients with different diseases, providing further information on key influencing variables and behavioral patterns is difficult. Therefore, nursing decision-makers may experience challenges in understanding the behavioral patterns of inpatients at different risk levels and providing appropriate personalized nursing services. This problem can be solved using machine learning methods such as support vector machine (SVM) and random forest (RF) [14, 15]. Notably, some machine learning methods have obvious advantages in predicting VTE risk. However, the main challenge is that the common machine learning methods for predicting VTE risk are black-box models, which lack interpretability, have poor transparency of results, and cannot be trusted by clinicians [16].

The present study aimed to further explore the key risk factors for VTE and the behavioral rules of different risk levels in urological inpatients by integrating the Caprini scale and an interpretable machine learning method. An interpretable prediction model was established.

Materials and methods

Research design

This study used data from the case hospital, cleaned the data, and transformed them into discrete data for analysis. Using the Boruta algorithm, 37 variables (conditional variables) of the Caprini scale were selected for feature selection. The feature variables that significantly contributed to the risk level (target variables) were retained. Regarding the sampling method, each data point had an approximately 80% probability of being sampled to the training set and a 20% probability of being sampled to the test set, which was executed five times. Furthermore, based on these training sets, a prediction model was established using the rough set (RS) method, and the corresponding decision rules were generated. The accuracy of the prediction model was verified using the test set. Finally, three commonly used machine learning methods, RF, SVM, and BPANN, were used to verify the accuracy of the model and were compared with the RS method. The design flow of this study is shown in Fig. 1.

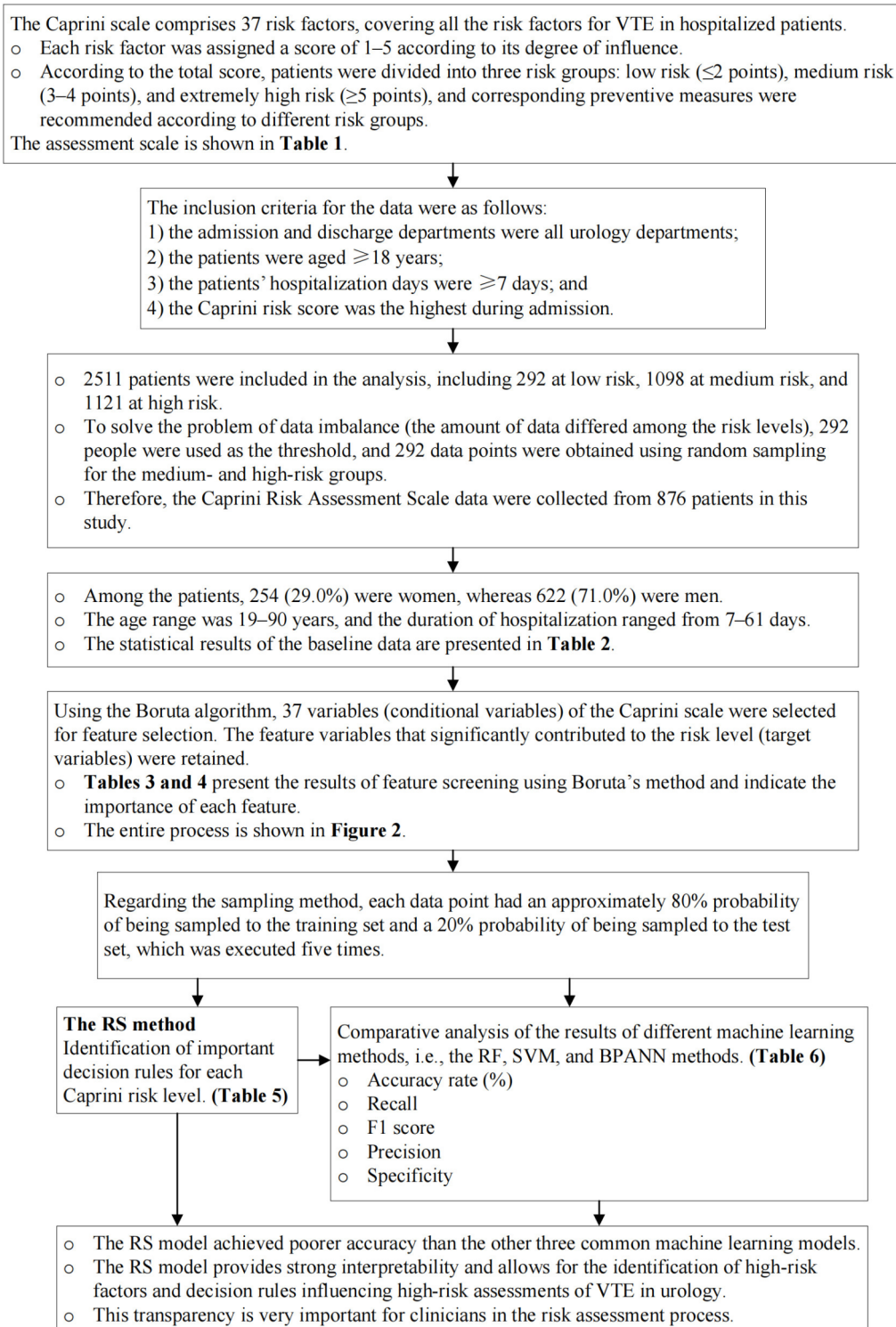


Fig. 1 The design flow of this study

Caprini scale

Since the late 1980s, Caprini's team has been conducting detailed individual risk assessments of medical and surgical patients [17]. The assessment scale uses a mixed

method based on evidence-based guidelines and consensus statements and further combines logic, emotion, and interviewer experience [18]. The assessment scale is called the Caprini scale. The Caprini scale comprises 37

risk factors, covering all the risk factors for VTE in hospitalized patients. Each risk factor was assigned a score of 1–5 according to its degree of influence. According to the total score, patients were divided into three risk groups: low risk (≤ 2 points), medium risk (3–4 points), and extremely high risk (≥ 5 points), and corresponding preventive measures were recommended according to different risk groups [19]. The scale has a good evaluation effect and has been used in many different patient groups, such as gynecologic [20], trauma [21], and pulmonary [22] patients. The assessment scale is shown in Table 1.

The Boruta method

Boruta is a feature-selection algorithm based on an RF classifier [23]. Boruta's algorithm randomizes the original feature variables, allowing each feature to create a corresponding "shadow" feature variable with a value obtained by rearranging the original feature values, classifying all the features of the extended variable, and calculating the importance of all the features [24]. The Boruta feature selection algorithm aims to select the feature set most relevant to the dependent variables. The Boruta algorithm retains a feature set that significantly contributes to the classification rather than the highest contribution, thus reducing overfitting [25]. The specific steps were as follows:

Step 1: When modeling for the first time, the original feature variables are copied as shadow variables, and each eigenvalue of the original feature matrix is randomized as the value of the shadow variables and spliced into a new feature matrix.

Step 2: Taking the new feature matrix as the input, the model of the important features is indicated as the output.

Step 3: The Z values of the original and shadow feature variables are calculated.

Step 4: The Z value of the shadow variable is indicated as Z_{max} , the original characteristic variable with a Z value greater than Z_{max} is indicated as "important," and the original characteristic variable with a Z value less than Z_{max} is indicated as "unimportant."

Step 5: Shadow variables are eliminated, and steps 1–4 are repeated until all variables are indicated as "important" or "unimportant."

The RS method

The RS method is based on set theory, a mathematical theory that deals mainly with qualitative or imprecise data, information, and knowledge [26]. Because of the large amount of incomplete and inconsistent information in the real world that cannot be analyzed using general statistical methods, the RS method can generalize data collection, determine the hidden data types and data

correlations, and produce usable classification rules [27]. The RS method basically aims to form concepts and rules through relational database classification induction and realize knowledge discovery through the classification of equivalence relations and the approximation of goals [28]. To date, the RS method has been applied to various topics, such as pressure injury risk [27], medical service quality [29], patient diagnosis [30], and lung cancer diagnosis [31].

Ethics approval

This study was approved by the Ethics Committee of the Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University (approval number: K20230202) and informed consent was waived. At the same time, the study complied with the principles of the Helsinki Declaration. The data were obtained from the VTE information system of the case hospital, and an information confidentiality agreement was signed with the hospital. During information extraction, only the hospitalization number of patients was kept, and their personal information was not collected, except for age. Throughout the research process, the privacy and confidentiality of the patients were protected.

Data collection and participants

Caprini risk assessment data were obtained from the urology department of a general hospital in Zhejiang Province, China. The data were collected between December 2019 and July 2022. The inclusion criteria for the data were as follows: (1) the admission and discharge departments were all urology departments; (2) the patients were aged ≥ 18 years; (3) the patients' hospitalization days were ≥ 7 days; and (4) the Caprini risk score was the highest during admission. Therefore, 2511 patients were included in the analysis, including 292 at low risk, 1098 at medium risk, and 1121 at high risk. To solve the problem of data imbalance (the amount of data differed among the risk levels), 292 people were used as the threshold, and 292 data points were obtained using random sampling for the medium- and high-risk groups. Therefore, the Caprini Risk Assessment Scale data were collected from 876 patients in this study. Among the patients, 254 (29.0%) were women, whereas 622 (71.0%) were men. The age range was 19–90 years, and the duration of hospitalization ranged from 7 to 61 days. The statistical results of the baseline data are presented in Table 2.

Results

Identification of important features affecting Caprini risk levels using Boruta's method

Tables 3 and 4 present the results of feature screening using Boruta's method and indicate the importance of

Table 1 The Caprini risk assessment model

Attributes	Value
Condition attributes	
Age (C ₁)	Age less than 40 years = 0; Age 41–60 years = 1; Age 61–74 years = 2; Age over 75 years = 3
Minor Surgery planned (C ₂)	Yes = 1; No = 0
Obesity (BMI > 25) (C ₃)	Yes = 1; No = 0
Pregnancy or postpartum (< 1 month) (C ₄)	Yes = 1; No = 0
History of unexplained stillborn infant recurrent spontaneous abortion (> 3) premature birth with toxemia or growth restricted infant (C ₅)	Yes = 1; No = 0
Oral contraceptives or hormone replacement therapy (C ₆)	Yes = 1; No = 0
Swollen legs (current) (C ₇)	Yes = 1; No = 0
Varicose veins (C ₈)	Yes = 1; No = 0
Sepsis (< 1 month) (C ₉)	Yes = 1; No = 0
Serious lung disease incl. pneumonia (< 1 month) (C ₁₀)	Yes = 1; No = 0
Abnormal pulmonary function (COPD) (C ₁₁)	Yes = 1; No = 0
Acute myocardial infarction (C ₁₂)	Yes = 1; No = 0
Congestive heart failure (< 1 month) (C ₁₃)	Yes = 1; No = 0
History of inflammatory bowel disease (C ₁₄)	Yes = 1; No = 0
Medical patient currently at bed rest (C ₁₅)	Yes = 1; No = 0
Other risk (C ₁₆)	Yes = 1; No = 0
Arthroscopic surgery (C ₁₇)	Yes = 1; No = 0
Major surgery (> 45 min) (C ₁₈)	Yes = 1; No = 0
Laparoscopic surgery (> 45 min) (C ₁₉)	Yes = 1; No = 0
Patient confined to bed (> 72 h) (C ₂₀)	Yes = 1; No = 0
Malignancy (present or previous) (C ₂₁)	Yes = 1; No = 0
Immobilizing plaster cast (< 1 month) (C ₂₂)	Yes = 1; No = 0
Central venous access (C ₂₃)	Yes = 1; No = 0
Family history of thrombosis (C ₂₄)	Yes = 1; No = 0
Positive Factor V Leiden (C ₂₅)	Yes = 1; No = 0
Positive Prothrombin 20,210 A (C ₂₆)	Yes = 1; No = 0
Positive lupus anticoagulant (C ₂₇)	Yes = 1; No = 0
Elevated anticardiolipin antibodies (C ₂₈)	Yes = 1; No = 0
Elevated serum homocysteine (C ₂₉)	Yes = 1; No = 0
Heparin-induced thrombocytopenia (HIT) (C ₃₀)	Yes = 1; No = 0
History of DVT/PE (C ₃₁)	Yes = 1; No = 0
Other congenital or acquired thrombophilia (C ₃₂)	Yes = 1; No = 0
Elective major lower extremity arthroplasty (C ₃₃)	Yes = 1; No = 0
Stroke (< 1 month) (C ₃₄)	Yes = 1; No = 0
Hip, pelvis or leg fracture (< 1 month) (C ₃₅)	Yes = 1; No = 0
Acute spinal cord injury (paralysis) (< 1 month) (C ₃₆)	Yes = 1; No = 0
Multiple trauma (< 1 month) (C ₃₇)	Yes = 1; No = 0
Decision attribute	
Risk level (D)	Low Risk (Score: less than or equal to 2) = 1; Moderate risk (Score: 2–4) = 2; High Risk (Score: greater than 5) = 3

Note: C: Condition attributes, D: Decision attribute

each feature. Table 3 presents the initial modeling results, with all features divided into three categories: confirmed, rejected, and tentative. Table 4 presents the results of the final modeling. All features were classified as “confirmed”

or “rejected.” “Confirmed” implied that the feature was important to the entire model or had a significant influence, and this feature was retained in the model. Rejected implied that the feature should be excluded. “Tentative”

Table 2 The background description

Items	n (%)
Total	n=876
Sex	
Male	622 (71.0%)
Female	254 (29.0%)
Age (years old)	
≤ 40	93(10.6%)
41–60	319 (36.4%)
61–74	339(38.7%)
≥ 75	125 (14.3%)
Average length of stay (days)	
7–10	322 (36.8%)
11–20	430 (49.1%)
21–30	100(11.4%)
>30	24 (2.7%)

implied that the feature could not be retained or excluded presently and required further processing.

According to the first modeling results in Table 3, “(C₁),” “(C₂),” “(C₃),” “(C₈),” “(C₉),” “(C₁₀),” “(C₁₁),” “(C₁₆),” “(C₁₈),” “(C₁₉),” “(C₂₀),” “(C₂₁),” “(C₂₃),” “(C₃₁),” and “(C₃₄)” were classified as “confirmed,” “(C₃₂)” was classified as “tentative,” and “(C₄),” “(C₅),” “(C₆),” “(C₇),” “(C₁₂),” “(C₁₃),” “(C₁₄),” “(C₁₅),” “(C₁₇),” “(C₂₂),” “(C₂₄),” “(C₂₅),” “(C₂₇),” “(C₂₈),” “(C₂₉),” “(C₃₀),” “(C₃₃),” “(C₃₅),” “(C₃₆),” and “(C₃₇)” were classified as “rejected.” After further analysis (Table 4), based on the median attribute score with the median score of the most important shadow attribute, “(C₃₂)” was confirmed to be important. The entire process is shown in Fig. 2.

Identification of important decision rules for each Caprini risk level using the RS method

Using the RS method, the data were modeled stochastically five times, and the low-risk group (D=1) produced 33 decision rules, among which five were indicated as important, affecting the low-risk assessment results of the Caprini risk score. The mid-risk group (D=2) produced 142 decision rules, among which three were indicated as important decision rules, which influenced the mid-risk assessment results in the Caprini risk score. The high-risk group (D=3) produced 189 decision rules, among which four were indicated as important decision rules, which influenced the high-risk assessment results of the Caprini risk score. The important decision rules for the three risk levels are presented in Table 5. The results of all decision rules are presented in the Appendix.

Among the most important decision rules for all risk levels, different conditional attributes/feature variables play a decisive role, indicating the circumstances under which the corresponding decision behavior will occur. Among the five most important decision rules of the low-risk group, nine conditional attributes were identified:

“(C₁₈),” “(C₁₆),” “(C₂₁),” “(C₂₀),” “(C₁₀),” “(C₁₉),” “(C₉),” “(C₁),” and “(C₃).” Among the three most important decision rules at the mid-risk level, 11 conditional attributes/feature variables were identified: “(C₃₁),” “(C₁₁),” “(C₂₀),” “(C₁₀),” “(C₁₆),” “(C₂₁),” “(C₁₈),” “(C₁),” “(C₂),” “(C₃),” and “(C₃₄).” Among the four decision rules of the high-risk level, four conditional attributes were identified: “(C₂₁),” “(C₁),” “(C₁₈),” and “(C₁₉).”

Results of the five random validations

Regarding the sampling methods, each dataset had an 80% probability of being sampled to the training set and a 20% probability of being sampled to the test set. Therefore, the amount of data in the training and test sets of each model changed differently to increase the amount of data in different modeling and testing situations. Based on the amount of data, RS, RF, SVM, and BPANN models were used to establish a prediction model. On the basis of this model, the accuracy, F1-score, recall (sensitivity), specificity and precision were calculated. The results are shown in Table 6. Table 6 presents the average accuracy, F1-score, recall (sensitivity), specificity and precision of each method. The average accuracy rates of the RS, RF, SVM, and BPANN algorithms were 79.5%, 87.9%, 92.6%, and 97.2%, respectively. The average F1-scores of the RS, RF, SVM and BPANN models were 0.815, 0.900, 0.937, and 0.974, respectively. The average recalls were 0.902, 0.915, 0.926, and 0.971, respectively. The average specificity was 0.900, 0.946, 0.978, and 0.971, respectively. The average precision values were 0.754, 0.874, 0.954, and 0.975, respectively.

Discussions

Critical risk attributes associated with important decision rules

The Caprini score is a widely verified VTE risk assessment model with high reliability that is widely used in various surgical specialties and is not limited to urology. There were 37 variables in the Caprini scale; however, some were not associated with urological diseases. Machine learning was used to screen variables, reduce the number of variables in the Caprini scale, and focus on variables highly correlated with urological diseases, thus guiding urologists to predict VTE risk more efficiently and accurately.

The Caprini scale is divided into three risk levels, each of which has different important decision rules. At the low-risk level, among the five important decision rules, there were nine risk attributes: “(C₁₈) (five times),” “(C₁₆) (four times),” “(C₂₁) (five times),” “(C₁₉) (five times),” “(C₉) (five times),” “(C₁) (five times),” “(C₃) (four times),” “(C₁₀) (three times),” and “(C₂₀) (three times).” At the mid-risk level, among the three important decision rules, there were 11 risk attributes: “(C₃₁) (three times),” “(C₁₁) (three

Table 3 Results of preliminary estimation and screening

	meanImp	medianImp	minImp	maxImp	normHits	decision
C ₁	71.64265	73.00339	44.66874	80.90723	1	Confirmed
C ₂	30.64722	31.19163	18.8274	35.53343	1	Confirmed
C ₃	45.89759	46.55414	28.14239	53.86793	1	Confirmed
C ₄	-1.36835	-1.48543	-3.28782	0.364843	0	Rejected
C ₅	0	0	0	0	0	Rejected
C ₆	0	0	0	0	0	Rejected
C ₇	-1.43638	-1.55038	-2.97635	0.120213	0	Rejected
C ₈	18.3633	18.35171	16.18967	20.34223	1	Confirmed
C ₉	18.62949	18.79633	12.30014	22.15236	1	Confirmed
C ₁₀	32.21408	32.51045	25.76599	35.09444	1	Confirmed
C ₁₁	10.39494	10.45248	7.647085	13.02257	0.984925	Confirmed
C ₁₂	0	0	0	0	0	Rejected
C ₁₃	0	0	0	0	0	Rejected
C ₁₄	0	0	0	0	0	Rejected
C ₁₅	-0.64248	-0.55464	-1.98551	1.109898	0	Rejected
C ₁₆	26.84088	27.15344	19.295	30.51425	1	Confirmed
C ₁₇	0	0	0	0	0	Rejected
C ₁₈	78.386	80.12163	46.42951	88.30215	1	Confirmed
C ₁₉	48.52522	49.32282	33.01249	54.96861	1	Confirmed
C ₂₀	18.35356	18.39682	15.16165	21.02918	1	Confirmed
C ₂₁	78.53399	80.13586	50.18942	88.82413	1	Confirmed
C ₂₂	0	0	0	0	0	Rejected
C ₂₃	17.33234	17.33434	15.3397	19.23019	1	Confirmed
C ₂₄	0	0	0	0	0	Rejected
C ₂₅	0	0	0	0	0	Rejected
C ₂₆	0	0	0	0	0	Rejected
C ₂₇	0	0	0	0	0	Rejected
C ₂₈	0	0	0	0	0	Rejected
C ₂₉	0	0	0	0	0	Rejected
C ₃₀	0	0	0	0	0	Rejected
C ₃₁	16.61551	16.67109	14.44252	18.85935	1	Confirmed
C ₃₂	4.048248	4.002713	2.553681	6.060474	0.517588	Tentative
C ₃₃	0	0	0	0	0	Rejected
C ₃₄	24.99644	25.02482	21.13339	27.94297	1	Confirmed
C ₃₅	0	0	0	0	0	Rejected
C ₃₆	0	0	0	0	0	Rejected
C ₃₇	0	0	0	0	0	Rejected

times), “(C₁₀) (three times),” “(C₁₆) (three times),” “(C₂₁) (three times),” “(C₂) (three times),” “(C₁₈) (three times),” “(C₁) (three times),” “(C₃) (three times),” “(C₉) (two times),” and “(C₃₄) (two times)”. For the high-risk level, among the four important decision rules, there were four risk attributes: “(C₂₁) (four times),” “(C₁) (three times),” “(C₁₈) (two times),” and “(C₁₉) (seven times).” Based on the above results, “(C₁) (three times),” “(C₁₈) (two times),” and “(C₂₁) (seven times)” were critical risk attributes for VTE risk stratification when the Caprini score was used for VTE risk assessment of urological inpatients. In addition to these three critical risk attributes, obesity, history of VTE, and lung disease were also important factors. Previous studies have confirmed that these factors have an

important impact on VTE risk classification. Therefore, we discuss this according to the following five factors:

Age

Age is a key characteristic variable for VTE risk. In a study on risk factors for VTE in different majors and directions, it was proven that age is an important factor influencing VTE risk classification. For people aged >40 years, VTE risk gradually increases with age and doubles every 10 years [32]. One study showed that among VTE events in the community, the proportion of people aged >65 years was as high as 60% [33].

In addition, with increasing age, other factors affecting VTE risk will also change, especially in older patients aged >65 years; with organ aging, the pathophysiology

Table 4 Results of final estimation and screening

	meanImp	medianImp	minImp	maxImp	normHits	decision
C ₁	71.64265	73.00339	44.66874	80.90723	1	Confirmed
C ₂	30.64722	31.19163	18.8274	35.53343	1	Confirmed
C ₃	45.89759	46.55414	28.14239	53.86793	1	Confirmed
C ₄	-1.36835	-1.48543	-3.28782	0.364843	0	Rejected
C ₅	0	0	0	0	0	Rejected
C ₆	0	0	0	0	0	Rejected
C ₇	-1.43638	-1.55038	-2.97635	0.120213	0	Rejected
C ₈	18.3633	18.35171	16.18967	20.34223	1	Confirmed
C ₉	18.62949	18.79633	12.30014	22.15236	1	Confirmed
C ₁₀	32.21408	32.51045	25.76599	35.09444	1	Confirmed
C ₁₁	10.39494	10.45248	7.647085	13.02257	0.984925	Confirmed
C ₁₂	0	0	0	0	0	Rejected
C ₁₃	0	0	0	0	0	Rejected
C ₁₄	0	0	0	0	0	Rejected
C ₁₅	-0.64248	-0.55464	-1.98551	1.109898	0	Rejected
C ₁₆	26.84088	27.15344	19.295	30.51425	1	Confirmed
C ₁₇	0	0	0	0	0	Rejected
C ₁₈	78.386	80.12163	46.42951	88.30215	1	Confirmed
C ₁₉	48.52522	49.32282	33.01249	54.96861	1	Confirmed
C ₂₀	18.35356	18.39682	15.16165	21.02918	1	Confirmed
C ₂₁	78.53399	80.13586	50.18942	88.82413	1	Confirmed
C ₂₂	0	0	0	0	0	Rejected
C ₂₃	17.33234	17.33434	15.3397	19.23019	1	Confirmed
C ₂₄	0	0	0	0	0	Rejected
C ₂₅	0	0	0	0	0	Rejected
C ₂₆	0	0	0	0	0	Rejected
C ₂₇	0	0	0	0	0	Rejected
C ₂₈	0	0	0	0	0	Rejected
C ₂₉	0	0	0	0	0	Rejected
C ₃₀	0	0	0	0	0	Rejected
C ₃₁	16.61551	16.67109	14.44252	18.85935	1	Confirmed
C ₃₂	4.048248	4.002713	2.553681	6.060474	0.517588	Confirmed
C ₃₃	0	0	0	0	0	Rejected
C ₃₄	24.99644	25.02482	21.13339	27.94297	1	Confirmed
C ₃₅	0	0	0	0	0	Rejected
C ₃₆	0	0	0	0	0	Rejected
C ₃₇	0	0	0	0	0	Rejected

of the coagulation system changes, and the incidence of chronic diseases affecting VTE risk, such as cardiovascular, cerebrovascular, and lung diseases, will also increase significantly. Therefore, age is an important factor influencing VTE risk, with VTE risk increasing significantly with increasing age. The influence of age on VTE risk can be mainly explained by two factors: traditional and unconventional risk factors.

Traditional risk factors include the following: (1) Immobility or decreased activity: With a gradual increase in age, the amount of activity tends to decrease, and the long-term immobility of older patients leads to an increase in blood viscosity, which is also an important factor influencing VTE risk. In addition, older patients have an increased risk of stroke and fracture, which can

increase the probability of bed rest. Notably, VTE risk is highest in the first 4 weeks of bed rest. (2) Increases in complications: With increasing age, the incidence of diseases related to VTE, such as cancer, heart failure, stroke, and diabetes, increases significantly.

Unconventional risk factors include reduced muscle strength and venous insufficiency. Muscle strength begins to decline from the age of 50–55 years [16], with degenerative changes in the lower limb joints, leading to the loss of leg muscles and a decrease in nerve regulation function, further leading to a decrease in muscle strength and seriously affecting the blood pump function of calf muscles. A review showed that thrombosis in older patients is associated with blood stasis and reflux caused by venous dysfunction, which may be caused by

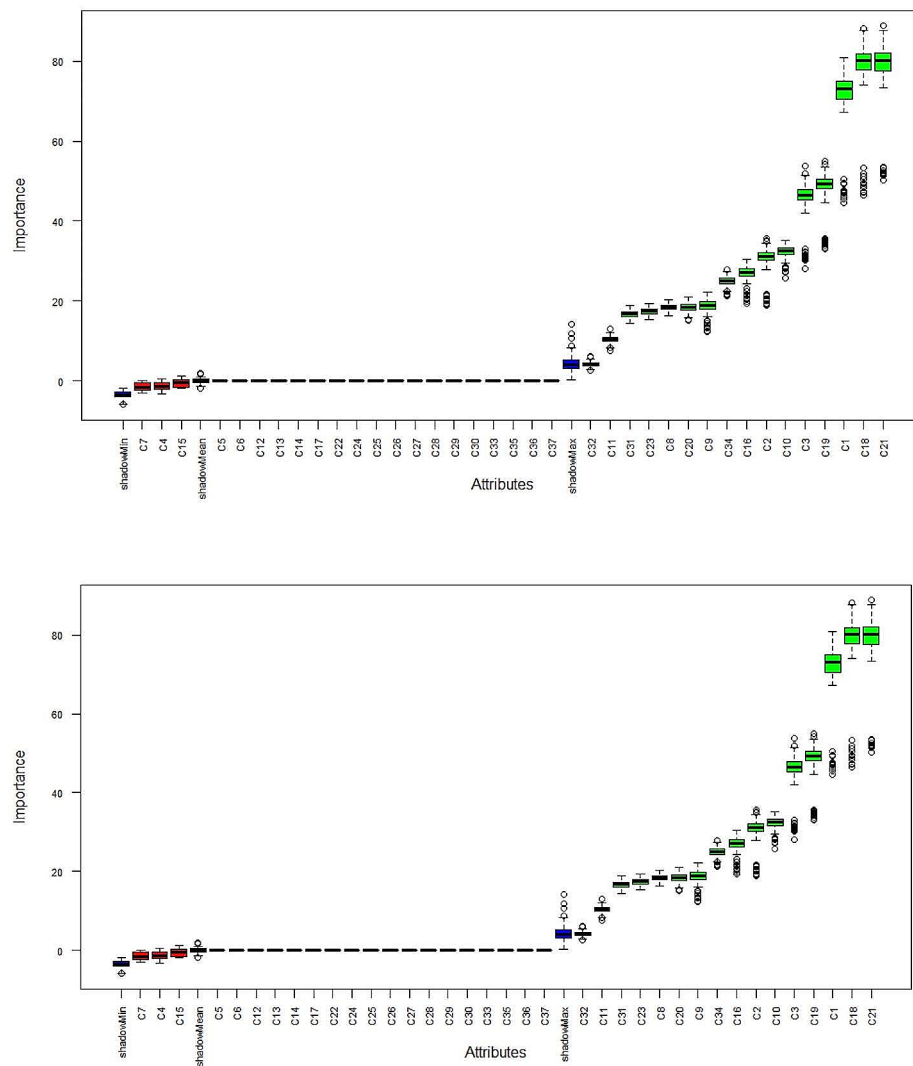


Fig. 2 The importance of each item/conditional attribute

the decreased blood pumping function of the calf muscles [34].

Obesity (body mass index)

“Obesity (body mass index [BMI]>25 kg/m²; C₃)” is a high-risk factor for VTE [35]. An increase in BMI is directly proportional to VTE risk. In one study, VTE risk in obese people with a BMI≥30 kg/m² increased two-fold compared with that in the general population [36]. Obesity is an important risk factor for myocardial infarction [35]. Considering VTE risk, obesity and myocardial infarction are associated with hypersuperposition. A study showed that the incidence of VTE in patients exposed to two factors increased threefold compared with that in patients exposed to a single factor, and the comprehensive effect of the two exposures exceeded the sum of the individual effects [35, 37, 38]; obesity is a

prerequisite for the two factors. Obesity is often accompanied by a hypercoagulable and inflammatory state in the blood, and a hypercoagulable state is an important condition for VTE. A meta-analysis of the correlation between C-reactive protein levels and VTE risk showed that the inflammatory state was positively correlated with VTE risk [39]. Therefore, similar to age, obesity is an important variable associated with VTE risk in urological patients. Based on this factor, medical staff can quickly and conveniently obtain corresponding information, perform a rough analysis of VTE risk in patients, and increase attention to VTE risk.

Surgical factors

In the present study, the characteristic variables associated with surgery included “(C₂)”, “(C₁₈)”, and “(C₁₉).” Surgery is the main treatment method used in urology;

Table 5 The most important decision rules in 5-times random modelling

No.	Condition (IF)	Decision (Then)	Support size	Laplace
1	IF [C21 is 0] and [C16 is 0] and [C20 is 0] and [C10 is 0] and [C9 is 0] and [C19 is 0] and [C18 is 0] and [C1 is 1] and [C3 is 0]	Low Risk (1)	78	0.9753
2	IF [C18 is 0] and [C16 is 0] and [C21 is 0] and [C10 is 0] and [C19 is 0] and [C9 is 0] and [C1 is 1] and [C3 is 0]	Low Risk (1)	75	0.9744
3	IF [C18 is 0] and [C16 is 0] and [C21 is 0] and [C20 is 0] and [C10 is 0] and [C19 is 0] and [C9 is 0] and [C1 is 1] and [C3 is 0]	Low Risk (1)	73	0.9737
4	IF [C21 is 0] and [C20 is 0] and [C19 is 0] and [C9 is 0] and [C18 is 0] and [C1 is 1] and [C3 is 0]	Low Risk (1)	73	0.9737
5	IF [C18 is 0] and [C16 is 0] and [C21 is 0] and [C19 is 0] and [C9 is 0] and [C2 is 0] and [C1 is 1]	Low Risk (1)	72	0.9733
6	IF [C34 is 0] and [C31 is 0] and [C11 is 0] and [C20 is 0] and [C16 is 0] and [C10 is 0] and [C21 is 0] and [C2 is 0] and [C18 is 1] and [C1 is 2] and [C3 is 0]	Middle Risk (2)	42	0.9556
7	IF [C31 is 0] and [C11 is 0] and [C20 is 0] and [C16 is 0] and [C10 is 0] and [C21 is 0] and [C2 is 0] and [C18 is 1] and [C1 is 2] and [C3 is 0]	Middle Risk (2)	42	0.9566
8	IF [C34 is 0] and [C31 is 0] and [C11 is 0] and [C16 is 0] and [C10 is 0] and [C21 is 0] and [C2 is 0] and [C18 is 1] and [C1 is 2] and [C3 is 0]	Middle Risk (2)	38	0.9512
9	IF [C18 is 1] and [C1 is 3]	High Risk (3)	48	0.9608
10	IF [C21 is 1] and [C18 is 1]	High Risk (3)	45	0.9583
13	IF [C21 is 1] and [C1 is 3]	High Risk (3)	39	0.9524
14	IF [C21 is 1] and [C19 is 1]	High Risk (3)	39	0.9524

Support size: The number of samples that support this rule

Laplace: The larger the number, the higher the laplace

however, surgery is recognized as the main risk factor for VTE, and the incidence of VTE increases significantly after major abdominal and pelvic surgeries [40]. From a pathophysiological perspective, surgery-induced vascular injury can easily lead to platelet aggregation and fibrosis repair induced by anticoagulant factors [41].

In addition, surgery is associated with risk factors for VTE, such as braking, hypercoagulability, and an inflammatory state. For new minimally invasive surgical methods, such as laparoscopy and robotic surgery, long-term pneumoperitoneum and positioning compress the main veins in the abdominal cavity [42], aggravating venous blood stasis and overlapping with other risk factors.

The incidence of VTE differs among different urological surgical types. A study on the incidence of DVT and its influencing factors in urology showed that prostatectomy (including traditional laparotomy and transurethral resection) exhibited the highest incidence, followed by cystectomy and urinary calculus surgery. Prostatectomies and cystectomies are complicated pelvic surgeries performed in urology [43].

Malignant tumors

In the present study, “(C₂₁)” was a risk factor for VTE, and many studies have confirmed that malignant tumors are an independent risk factor for VTE [39, 44]. According to relevant research, the incidence of VTE in patients with malignant tumors is 4–5 times greater than that in patients with nonmalignant tumors [45]. Approximately 20% of patients with VTE have malignant tumors, which are a critical cause of death in patients with VTE. At the pathophysiological level, malignant tumors, as exogenous

factors, activate coagulation factor X and promote platelet activation and fibrin synthesis [46]. The risk factors for malignant tumors are also associated with tumor stage and treatment methods such as surgery, central venous catheter placement or infusion port placement, and chemotherapy (a risk factor for venous endothelial injury). Each factor is an independent risk factor for VTE that produces a risk superposition effect [45, 46]. In urology, malignant tumors in the bladder, prostate, and kidney are common. Treatment methods for malignant tumors in the urinary system mainly include surgery, chemotherapy, and central venous catheterization. In addition, long-term catheter and central venous catheter indwelling increases the risk of catheter-related infection, which indirectly leads to inflammation-related VTE risk, and the comprehensive risk is far greater than the cumulative sum of the individual risks. In addition, weakness, immobility (including postoperative immobilization), and pain (including cancer and postoperative pain) caused by malignant tumors increase the risk of VTE [44].

Lung disease

In the present study, the important factors associated with lung diseases were “(C₁₀)” and “(C₁₁)”, which mainly included abnormal lung function, chronic obstructive pulmonary disease (COPD), and other lung diseases. According to a previous study on the correlation between COPD and VTE, COPD is an independent risk factor for VTE. Pulmonary embolism (PE) is the primary manifestation of VTE in patients with COPD compared with other patients. The incidence of cerebral vein thrombosis is lower in patients with COPD than in patients with PE

Table 6 The prediction model quality scores of each method randomly executed 5 times

	RSA	RF	SVM	BPANN
	Accuracy rate (%)			
1	79.4%	89.7%	94.9%	97.7%
2	79.1%	86.6%	93.6%	96.0%
3	73.7%	87.4%	94.0%	98.3%
4	82.6%	89.3%	87.4%	97.2%
5	82.9%	86.7%	93.0%	96.6%
Average	79.5%	87.9%	92.6%	97.2%
	Recall (sensitivity)			
1	0.821	0.943	0.975	0.965
2	0.908	0.866	0.944	0.964
3	0.889	0.902	0.946	0.976
4	0.938	0.933	0.814	0.977
5	0.954	0.933	0.952	0.965
Average	0.902	0.915	0.926	0.971
	F1-score			
1	0.807	0.925	0.969	0.964
2	0.797	0.889	0.941	0.980
3	0.821	0.888	0.948	0.979
4	0.834	0.904	0.881	0.980
5	0.802	0.894	0.949	0.970
Average	0.815	0.900	0.937	0.974
	Precision			
1	0.824	0.909	0.965	0.967
2	0.715	0.810	0.939	0.980
3	0.780	0.886	0.952	0.970
4	0.760	0.892	0.978	0.982
5	0.692	0.873	0.938	0.978
Average	0.754	0.874	0.954	0.975
	Specificity			
1	0.900	0.947	0.977	0.982
2	0.897	0.932	0.968	0.913
3	0.871	0.937	0.971	0.991
4	0.915	0.950	0.977	0.988
5	0.917	0.964	0.995	0.983
Average	0.900	0.946	0.978	0.971

[47], and the recurrence, bleeding, and death risks associated with VTE in patients with COPD are greater than those in patients without COPD [47–49].

Application of a practical prevention strategy based on important decision rules

Combining VTE risk assessment results with clinical preventive measures is important in VTE risk assessment. In the present study, the Caprini assessment scale was used to assess the risk of VTE in urological inpatients. There were four decision rules with support sizes > 35 in the high-risk group and three with support sizes > 35 in the medium-risk group. Clinicians, nurses, and medical administrators should pay attention to the decision-making information contained in these rules and implement

appropriate preventive measures. The three risk levels are as follows:

High VTE risk decision rules

There were four high-risk decision rules for VTE: IF [C_{18} is 1] and [C_1 is 3], IF [C_{21} is 1] and [C_{18} is 1], IF [C_{21} is 1] and [C_1 is 3], and IF [C_{21} is 1] and [C_{19} is 1]. Based on these four decision rules, it can be concluded that the characteristics of patients with a high risk of VTE in urology include older patients aged ≥ 75 years who underwent large-scale open or tumor surgery and patients who underwent large-scale tumor or laparoscopic surgery. With the continuous development of surgical techniques, instruments, and materials, large-scale open surgery in urology has gradually been replaced by endoscopy. Minimally invasive surgery and minimally invasive surgery performed through the natural lumen have become the mainstream surgical methods in urology. However, the degree of internal injury in complex endoscopic surgery in urology is still relatively large. The variable “(C_{19})” is rare; however, it is suggested that “(C_{18})” be treated equally to “(C_{19})” in practical work.

Patients with high VTE risk are subdivided into two categories according to bleeding risk: (i) high VTE and low bleeding risks and (ii) high VTE and bleeding risks.

According to relevant research reports, advanced age, malignant tumors, diabetes, and recent surgical history are high-risk factors for bleeding [1, 2, 41, 46] and VTE. Therefore, appropriate preventive measures should be taken according to the results of the bleeding evaluation. Clinicians should be aware of the standardized diagnosis and treatment of VTE, the indications and contraindications for anticoagulation and thrombolytic therapy, and the corresponding drugs (including unfractionated heparin, low-molecular-weight heparin, warfarin, and new oral anticoagulants) and treatment equipment.

For patients with high VTE and low bleeding risk, based on their conditions, medical staff should actively conduct basic preventive measures against VTE (including health education, avoiding dehydration and breaking, getting out of bed early, engaging in functional exercise, avoiding lower limb vascular puncture, and raising the lower limbs in a timely manner) to reduce some risk factors. In addition, based on basic preventive measures, drug use alone or drug use combined with mechanical prevention (including elastic socks, intermittent inflation and compression devices, and plantar venous pumps) should be adopted [7, 48].

Patients with high VTE and bleeding risks were treated with mechanical prevention methods, the course of prevention was generally 7–14 days postoperatively, and major tumor surgery was postponed until 28 days postoperatively [7]. Urologists should prioritize older patients undergoing tumor surgery in their departments.

If necessary, they should invite VTE multidisciplinary diagnosis and treatment teams or specialists to participate in preoperative discussions, perioperative treatment plan formulation, informed notification, and other work and incorporate multidisciplinary team opinions into the quality control index system for such patients. Urologists and nursing staff should closely observe the clinical manifestations of VTE, such as cough, hemoptysis, chest pain, and lower limb swelling.

Medium VTE risk decision rules

There were three high-risk decision rules for VTE. By combining these findings, we can conclude that the characteristics of patients at risk of VTE in the urology department was 1 for “(C₁₈)” and 2 for “(C₁)”, indicating that surgical patients aged 61–74 years were included. These patients mainly undergo bladder and prostate resection, urinary calculi, and other nonmalignant tumor operations [50–52]. In patients at risk of VTE, the risk of bleeding should also be assessed. Drug or mechanical prevention methods have been adopted for patients with VTE who are at low risk of bleeding. However, patients with VTE with a high risk of bleeding should adopt mechanical prevention methods, and the course of prevention is the same as that in patients with a high risk of VTE.

Comparative analysis of the results of different machine learning methods

In this study, regarding the sampling method, each data point had an 80% probability of being sampled to the training set and a 20% probability of being sampled to the test set. Therefore, the amount of data in the training and test sets of each model changed dynamically to increase the amount of data in different modeling and testing situations. Based on this amount of data, the RS, RE, SVM, and BPANN algorithms were used to establish a prediction model and calculate the accuracy. In terms of prediction accuracy, the RE, SVM, and BPANN models reached >85% accuracy, with the SVM model reaching 92.6% accuracy and the BPANN model reaching 97.2% accuracy, which is excellent. The accuracy of the RS method was close to 80%, indicating poorer accuracy than that of the other three machine learning methods.

Models used for medical evaluation and decision-making should be transparent and easy to use. Medical staff should be able to compare the results of the decision rules and key characteristic variables based on their knowledge. A high degree of transparency and interpretability may increase the trust of medical staff in machine learning for building models. Additionally, the RS method can reveal the laws and potential causal relationships underlying the data. Moreover, RE, SVM and BPANN are “black-box models,” and the internal calculation model

and decision rules cannot be explained, which is not easily accepted by clinical medical staff and managers. Such models cannot reveal the laws and potential causal relationships underlying the data. Therefore, as an interpretable machine learning method, the accuracy of the RS model is acceptable in the present study. The model may help urological medical staff explore the characteristics of patients at high risk of VTE, establish clear decision rules, quickly identify patients at high risk of VTE, conduct follow-up preventive measures accurately and in a timely manner, improve VTE evaluation accuracy, standardize prevention rates, reduce VTE incidence, and achieve standardized prevention and treatment.

Explainable machine learning models fall into two categories. The first category includes intrinsically interpretable machine learning models, such as logistic regression, decision trees, Bayesian models, and machine learning models based on decision rules. These models directly provide a certain degree of information through method features and are easily understood by decision makers. The second category is the postevent interpretation method of prediction models, which provides other supplementary explanatory information for most prediction models, such as the SHapley Additive exPlanation (SHAP) model.

Compared with non-rule-based classifiers, decision-rule-based methods all show a certain degree of performance degradation [53]. Clinical credibility and the application of prediction models largely rely on how well doctors understand and interpret models. Evaluation indicators include the accuracy of predictions and the complexity of interpreting the results. Sometimes, achieving high accuracy conflicts with the difficulty of explanation, necessitating a balance between the two.

In this study, RS is an interpretable machine learning method falling into the first category. RS resolves data ambiguities using set theory, providing clear decision rules.

The RS-based predictive model provides physicians with decision rules that improve model traceability and information content. The RS model has further advantages in that dominance-based decision rules condense a range of attribute values into each rule, thereby maximizing information density. Despite potential slight performance differences, the RS model is valued for its accessibility, simplicity, and ease of interpretation. Rule-based approaches have further benefits in that they clearly indicate the patient characteristics most relevant to VTE risk. The rules are simple and easily understood, particularly for high-risk patients, and can be enforced by medical personnel, improving the transparency and interpretability of the classification process and strengthening the accessibility of the model, thus enhancing its credibility.

Limitations

This study has several limitations. (i) The data were collected from inpatients in the Urology Department of a general hospital in Zhejiang Province from December 2019 to July 2022; this was a single-center retrospective study. Since single-center studies cannot represent the Chinese urology population, a multicenter study is needed to verify the generalizability of our findings.

Additionally, machine learning has several limitations. (ii) The accuracy of machine learning predictions depends on the data quality of the Caprini assessment. Laboratory examination data are relatively objective and have few influencing factors. However, observation and subjective evaluations by doctors, nurses or patient's self-reported medical history and family history may lead to deviations between clinical data and actual conditions. Variability in observation information among different personnel can affect prediction models, which rely on historical data [54]. The development of a multicenter risk prediction model using Caprini assessment data from various institutions may mitigate some limitations and enhance the advantages of machine learning in big data mining [55].

(iii) Artificial intelligence (AI) can uncover subtle patterns or relationships hidden in traditional VTE risk modeling, providing more effective decision-making information. Since large amounts of information are being condensed, performance degradation may still occur during model development. Even so, machine learning remains important for understanding and assessing real-world VTE risk. Improved data quality can further enhance predictive performance [56].

(iv) Medical ethics, laws and regulations limit the widespread adoption of AI decision-making models in clinical practice. Even so, machine learning-based predictive models may improve the efficiency of medical staff. The acceptance of AI by medical professionals is also critical for the successful application of predictive models in clinical practice [57].

Finally, (v) the decision rules of different VTE risk levels identified in this study require further development for better interpretation. Further research is needed to establish accurate preventive measures based on these decision rules.

Conclusions

In this study, 37 indexes of the Caprini rating scale were selected, and the Boruta and RF machine learning algorithms were used to analyze high-risk factors for VTE. A prediction model based on decision rules was further developed. Previous studies were mostly single-factor regression studies, which can only explain the relationship between a single risk factor and VTE but cannot provide clear decision rules. This study focused on

explaining the relationship between multiple influencing factors and VTE risk. Compared with single-factor studies, this study clearly reveals the patient characteristics and decision rules for different risks of VTE and improves the efficiency and quality of doctors' evaluation. The results showed that age (C_1), major surgery (>45 min; C_{18}), laparoscopic surgery (>45 min; C_{19}), and malignancy (C_{21}) were the most important factors affecting the VTE risk classification of urological patients. Medical staff in the field of urology should focus on evaluation and prevention. The evaluation of VTE risk factors in urologic patients may improve the accuracy of VTE risk assessment and the effective prevention of VTE. The RS model simplifies the dimensions and can be used to evaluate the risk of VTE in urological inpatients on a large scale as a first step towards improving hospital-acquired VTE. It can be applied to preventing and treating VTE in hospitals to provide suggestions for managers.

Acknowledgements

This study like to thank Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, for their great support in data collection, interpretation of results, and clinical care experience.

Author contributions

Chao Liu and Yen-Ching Chuang participated in the study and drafted the manuscript. Fengmin Cheng, Yanjun Jin and Wei-Ying Yang were involved in the study design and data collection. Yen-Ching Chuang compiled the data and calculated the results. Fengmin Cheng, Yanjun Jin and Wei-Ying Yang participated in the analysis and discussion of the results. Ching-Wen Chien, Yen-Ching Chuang, Wei-Ying Yang and Yanjun Jin conceived the design of this study and participated in the coordination and communication of the entire study. All authors read and approved the final manuscript.

Funding

This study received financial support from Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University (ID: 24EZB01); Taizhou Science and Technology Project (No.22ywb01); Zhangjiakou Science and Technology Project (No.2322060D).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University (approval number: K20230202) and informed consent was waived. At the same time, the study complied with the principles of the Helsinki Declaration. The data were obtained from the VTE information system of the case hospital, and an information confidentiality agreement was signed with the hospital. During information extraction, only the hospitalization number of patients was kept, and their personal information was not collected, except for age. Throughout the research process, the privacy and confidentiality of the patients were protected.

Competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author details

¹Medical Department, The Second Affiliated Hospital of Hebei North University, Zhangjiakou 075000, Hebei, China

²Institute for Hospital Management, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, Guangdong, China

³Nursing Department, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai 317000, Zhejiang, China

⁴Business College, Taizhou University, Taizhou 318000, Zhejiang, China

⁵Institute of Public Health & Emergency Management, Taizhou University, Taizhou 318000, Zhejiang, China

⁶Key Laboratory of evidence-based Radiology of Taizhou, Linhai 317000, Zhejiang, China

⁷Department of Urology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai 317000, Zhejiang, China

Received: 27 February 2024 / Accepted: 7 August 2024

Published online: 16 August 2024

References

- Scarpa RM, Carrieri G, Gussoni G, et al. Clinically overt venous thromboembolism after urologic cancer surgery: results from the @RISTOS Study. *Eur Urol*. 2007;51(1):130–6. <https://doi.org/10.1016/j.eururo.2006.07.014>.
- Pandor A, Tonkins M, Goodacre S, et al. Risk assessment models for venous thromboembolism in hospitalised adult patients: a systematic review. *BMJ Open*. 2021;11(7):e045672. <https://doi.org/10.1136/bmjopen-2020-045672>.
- Wang P, Wang Y, Yuan Z, et al. Venous thromboembolism risk assessment of surgical patients in Southwest China using real-world data: establishment and evaluation of an improved venous thromboembolism risk model. *BMC Med Inform Decis Mak*. 2022;22(1):59. <https://doi.org/10.1186/s12911-022-01795-9>.
- Frees SK, Aning J, Black P, et al. A prospective randomized pilot study evaluating an ERAS protocol versus a standard protocol for patients treated with radical cystectomy and urinary diversion for bladder cancer. *World J Urol*. 2017;36(2):215–20. <https://doi.org/10.1007/s00345-017-2109-2>.
- Tikkinen KAO, Craigie S, Agarwal A, et al. Procedure-specific risks of thrombosis and bleeding in urological non-cancer surgery: systematic review and meta-analysis. *Eur Urol*. 2018;73(2):236–41. <https://doi.org/10.1016/j.eururo.2017.02.025>.
- Monagle P, Cuello CA, Augustine C, et al. American Society of Hematology 2018 Guidelines for management of venous thromboembolism: treatment of pediatric venous thromboembolism. *Blood Adv*. 2018;2(22):3292–316. <https://doi.org/10.1182/bloodadvances.2018024786>.
- Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur Respir J*. 2019;54(3):1901647. <https://doi.org/10.1183/13993003.01647-2019>.
- Law Y, Chan YC, Cheng SWK. Epidemiological updates of venous thromboembolism in a Chinese population. *Asian J Surg*. 2018;41(2):176–82. <https://doi.org/10.1016/j.asjsur.2016.11.005>.
- Fernando SM, Tran A, Cheng W, et al. VTE prophylaxis in critically ill adults. *Chest*. 2022;161(2):418–28. <https://doi.org/10.1016/j.chest.2021.08.050>.
- Hussain MH, Kim S, Khan AA, et al. Analysis of readmissions due to VTE—using hospital data to improve VTE prophylaxis compliance: A quality improvement project. *Clin Appl Thromb/Hemost*. 2023;29:10760296231181916. <https://doi.org/10.1177/10760296231181916>.
- Golemi I, Salazar Adum JP, Tafur A, et al. Venous thromboembolism prophylaxis using the Caprini score. *Disease-a-Month*. 2019;65(8):249–98. <https://doi.org/10.1016/j.disamonth.2018.12.005>.
- Lobastov K, Urbanek T, Stepanov E, et al. The thresholds of Caprini score associated with increased risk of venous thromboembolism across different specialties: a systematic review. *Ann Surg*. 2023;277(6):929–37. <https://doi.org/10.1097/SLA.0000000000005843>.
- Frankel J, Belanger M, Tortora J, et al. Caprini score and surgical times linked to the risk for venous thromboembolism after robotic-assisted radical prostatectomy. *Türk Üroloji Dergisi/Turkish J Urol*. 2020;46(2):108–14. <https://doi.org/10.5152/tud.2019.19162>.
- Liu H, Yuan H, Wang Y, et al. Prediction of venous thromboembolism with machine learning techniques in young-middle-aged inpatients. *Sci Rep*. 2021;11(1):12868. <https://doi.org/10.1038/s41598-021-92287-9>.
- Lei H, Zhang M, Wu Z, et al. Development and validation of a risk prediction model for venous thromboembolism in lung cancer patients using machine learning. *Front Cardiovasc Med*. 2022;9:845210. <https://doi.org/10.3389/fcvm.2022.845210>.
- Caprini JA. Individual risk assessment is the best strategy for thromboembolic prophylaxis. *Dis Mon*. 2010;56(10):552–9. <https://doi.org/10.1016/j.disamonth.2010.06.007>.
- Caprini JA, Arcelus JI, Hasty JH, et al. Clinical assessment of venous thromboembolic risk in surgical patients. *Semin Thromb Hemost*. 1991;17 Suppl 3:304–12.
- Motykie GD, Caprini JA, Arcelus JI, et al. Risk factor assessment in the management of patients with suspected deep venous thrombosis. *Int Angiol*. 2000;19(1):47–51.
- Cronin M, Dengler N, Krauss ES, et al. Completion of the updated Caprini risk assessment model (2013 Version). *Clin Appl Thromb/Hemost*. 2019;25:1076029619838052. <https://doi.org/10.1177/1076029619838052>.
- Lewis GK, Spaulding AC, Brennan E, et al. Caprini assessment utilization and impact on patient safety in gynecologic surgery. *Arch Gynecol Obstet*. 2023;308(3):901–12. <https://doi.org/10.1007/s00404-023-07038-0>.
- Hazeltine MD, Scott EM, Dorfman JD. An abbreviated Caprini model for VTE risk assessment in trauma. *J Thromb Thrombolysis*. 2022;53(4):878–86. <https://doi.org/10.1007/s11239-021-02611-3>.
- Łukaszuk RF, Nycz KP, Plens K, et al. Caprini VTE computerized risk assessment improves the use of thromboprophylaxis in hospitalized patients with pulmonary disorders. *Adv Clin Exp Med*. 2022;31(3):261–266. <https://doi.org/10.17219/acem/115080>.
- Rudnicki MBKWR. Feature selection with the Boruta Package. *J Stat Softw*. 2010;36(11):0–0. <https://doi.org/10.18637/jss.v036.i11>.
- Kursa MB, Jankowski A, Rudnicki WR. Boruta – a system for feature selection. *Fundamenta Informaticae*. 2010;101(4):271–85. <https://doi.org/10.3233/fi-2010-288>.
- Handhika T, Murni M, Fahreza RM. Boruta algorithm: an alternative feature selection method in credit scoring model. *Nucleation Atmospheric Aerosols*. 2023;0(0094–243X):0–0. <https://doi.org/10.1063/5.0114178>.
- Wei W, Liang J. Information fusion in rough set theory: an overview. *Inform Fusion*. 2019;48:107–18. <https://doi.org/10.1016/j.inffus.2018.08.007>.
- Chuang YC, Miao T, Feng C, et al. Exploration of pressure injury risk in adult inpatients: An integrated Braden scale and rough set approach. *Intensive Crit Care Nurs*. 2024;80:103567. <https://doi.org/10.1016/j.iccn.2023.103567>.
- Chacón-Gómez F, Cornejo ME, Medina J, et al. Rough set decision algorithms for modeling with uncertainty. *J Comput Appl Math*. 2023;0(0377–0427):115413. <https://doi.org/10.1016/j.cam.2023.115413>.
- Du ML, Tung TH, Tao P, et al. Application of rough set theory to improve outpatient medical service quality in public hospitals based on the patient perspective. *Front Public Health*. 2021;9:739119. <https://doi.org/10.3389/fpubh.2021.739119>.
- Das S, Sil J. Managing boundary uncertainty in diagnosing the patients of rural area using fuzzy and rough set. *J Healthc Inform Res*. 2022;6(1):1–47. <https://doi.org/10.1007/s41666-021-00109-4>.
- Kumari N, Acharjya DP. A hybrid rough set shuffled frog leaping knowledge inference system for diagnosis of lung cancer disease. *Comput Biol Med*. 2023;155:106662. <https://doi.org/10.1016/j.compbiomed.2023.106662>.
- Engbers MJ, Van Hylckama Vlieg A, Rosendaal FR. Venous thrombosis in the elderly: incidence, risk factors and risk groups. *J Thromb Haemost*. 2010;8(10):2105–12. <https://doi.org/10.1111/j.1538-7836.2010.03986.x>.
- Spencer FA GJ, Lessard D. Venous thromboembolism in the elderly. A community-based perspective. *Thromb Haemost*. 2008;780–8.
- Zhang X, Cai Q, Wang X, et al. Current use of rivaroxaban in elderly patients with venous thromboembolism (VTE). *J Thromb Thrombolysis*. 2021;52(3):863–71. <https://doi.org/10.1007/s11239-021-02415-5>.
- Horvei LD, Brækkan SK, Mathiesen EB, et al. Obesity measures and risk of venous thromboembolism and myocardial infarction. *Eur J Epidemiol*. 2014;29(11):821–30. <https://doi.org/10.1007/s10654-014-9950-z>.
- Horvei LD, Grimnes G, Hindberg K, et al. C-reactive protein, obesity, and the risk of arterial and venous thrombosis. *J Thromb Haemost*. 2016;14(8):1561–71. <https://doi.org/10.1111/jth.13369>.
- Sejrup JK, Tøndel BG, Morelli VM, et al. Joint effect of myocardial infarction and obesity on the risk of venous thromboembolism: The Tromsø Study. *J Thromb Haemost*. 2022;20(10):2342–9. <https://doi.org/10.1111/jth.15812>.
- Ntinopoulou P, Ntinopoulou E, Papatheanasiou IV, et al. Obesity as a risk factor for venous thromboembolism recurrence: a systematic review. *Medicina*. 2022;58(9):1290. <https://doi.org/10.3390/medicina58091290>.
- Kunutsor SK, Seidu S, Blom AW, et al. Serum C-reactive protein increases the risk of venous thromboembolism: a prospective study and meta-analysis of published prospective evidence. *Eur J Epidemiol*. 2017;32(8):657–67. <https://doi.org/10.1007/s10654-017-0277-4>.

40. Felder S, Rasmussen MS, King R, et al. Prolonged thromboprophylaxis with low molecular weight heparin for abdominal or pelvic surgery. *Cochrane Database Syst Rev.* 2019;8(3):CD004318. <https://doi.org/10.1002/14651858.CD004318.pub5>.
41. Sagalovich D, Say R, Kaouk J, et al. The role of extended venous thromboembolism prophylaxis following urologic pelvic surgery. *Urologic Oncology: seminars and original investigations.* 2018;36(3):83–7. <https://doi.org/10.1016/j.urolonc.2017.12.010>.
42. Elsayed AS, Ozair S, Iqbal U, et al. Prevalence and predictors of venous thromboembolism after robot-assisted radical cystectomy. *Urol.* 2021;149:146–53. <https://doi.org/10.1016/j.jurology.2020.11.014>.
43. Tang G, Qi L, Sun Z, et al. Evaluation and analysis of incidence and risk factors of lower extremity venous thrombosis after urologic surgeries: a prospective two-center cohort study using LASSO-logistic regression. *Int J Surg.* 2021;89:105948. <https://doi.org/10.1016/j.ijsu.2021.105948>.
44. Theochari NA, Theochari CA, Kokkinidis DG, et al. Venous thromboembolism after esophagectomy for cancer: a systematic review of the literature to evaluate incidence, risk factors, and prophylaxis. *Surg Today.* 2021;52(2):171–81. <https://doi.org/10.1007/s00595-021-02260-2>.
45. Fernandes CJ, Morinaga LTK, Alves JL, et al. Cancer-associated thrombosis: the when, how and why. *Eur Resp Rev.* 2019;28(151):180119. <https://doi.org/10.1183/16000617.0119-2018>.
46. Lutsey PL, Zakai NA. Epidemiology and prevention of venous thromboembolism. *Nat Reviews Cardiol.* 2022;20(4):248–62. <https://doi.org/10.1038/s41569-022-00787-6>.
47. Bertoletti L, Quenet S, Mismetti P, et al. Clinical presentation and outcome of venous thromboembolism in COPD. *Eur Resp J.* 2012;39(4):862–868. <https://doi.org/10.1186/s12873-022-00736-z>.
48. Harenberg J, Verhamme P. The dangerous liaisons between chronic obstructive pulmonary disease and venous thromboembolism. *Thromb Haemost.* 2020;120(3):363–5. <https://doi.org/10.1055/s-0039-1701012>.
49. Dong W, Zhu Y, Du Y, et al. Association between features of COPD and risk of venous thromboembolism. *Clin Respir J.* 2019;13(8):499–504. <https://doi.org/10.1111/crj.13051>.
50. Yang Y, Li X, Zhai Z, et al. Identification of prophylaxis and treatment for hospitalized patients associated with venous thromboembolism. *Chin Med J.* 2023;136(9):1111–3. <https://doi.org/10.1097/cm9.0000000000002237>.
51. Zeng H, Gao M, Chen J, et al. Incidence and risk factors of venous thromboembolism after percutaneous nephrolithotomy: a single-center experience. *World J Urol.* 2021;39(9):3571–7. <https://doi.org/10.1007/s00345-021-03658-w>.
52. Li K, Yu M, Li H, et al. Establishment of prediction models for venous thromboembolism in non-oncological urological inpatients – a single-center experience. *Int J General Med.* 2022;15:3315–24. <https://doi.org/10.2147/ijgm.S354288>.
53. Gil-Herrera E, Aden-Buie G, Yalcin A, et al. Rough set theory based prognostic classification models for hospice referral. *BMC Med Inform Decis Mak.* 201;15(1):98. <https://doi.org/10.1186/s12911-015-0216-9>.
54. Chiasakul T, Lam BD, McNichol M, et al. Artificial intelligence in the prediction of venous thromboembolism: a systematic review and pooled analysis. *Eur J Haematol.* 2023;111(6):951–62. <https://doi.org/10.1111/ejh.14110>.
55. Hassan AM, Rajesh A, Asaad M, et al. Artificial intelligence and machine learning in prediction of surgical complications: current state, applications, and implications. *Am Surg.* 2023;89(1):25–30. <https://doi.org/10.1177/00031348221101488>.
56. He L, Luo L, Hou X, et al. Predicting venous thromboembolism in hospitalized trauma patients: a combination of the Caprini score and data-driven machine learning model. *BMC Emerg Med.* 2021;21(1):60. <https://doi.org/10.1186/s12873-021-00447-x>.
57. Arina P, Kaczorek MR, Hofmaenner DA, et al. Prediction of complications and prognostication in perioperative medicine: a systematic review and PROBAST assessment of machine learning tools. *Anesthesiology.* 2024;140(1):85–101. <https://doi.org/10.1097/aln.0000000000004764>.

Publisher's Note

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