Prediction of the Risk of Preoperative Deep Vein Thrombosis in Non-fractured Patients Awaiting Total Joint Arthroplasty Using a Support Vector Machine

Haosheng Wang
Taizhou Central Hospital
Tingting Fan
Baoji City Hospital of Traditional Chinese Medicine
Yanhua Chen
Wuhan University State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing
Wenle Li
Xianyang Central Hospital
Fujiang Zhao
Taizhou Central Hospital
Qiuming Liu
Taizhou Central Hospital
Mingyu Yang (616353108@qq.com)
Taizhou Central Hospital

Research Article

Keywords: Risk factors, Outcomes, Support vector machine, Total joint arthroplasty, Deep vein thrombosis

Posted Date: August 31st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-844532/v1

Abstract

Background: We developed a potential useful alternative prediction model based on the support vector machine (SAM) algorithm to predict the risk of preoperative deep vein thrombosis (DVT) in non-fractured patients awaiting total joint arthroplasty (TJA).

Methods: From March 2015 to August 2020, a retrospective review of the preoperative ultrasound examination findings of lower extremity venous vessels was performed on non-fractured patients of 369 elective TJA. Based on the ultrasound examination findings of preoperative lower extremely venous vessels, these patients were divided into two groups: the DVT group and the Non-DVT group. We collected the clinical, imaging, and laboratory findings from an electronic medical record system. These variables were imported into univariate, multivariate and logistic regression analysis to identify the risk factor for preoperative DVT. According to published literature and clinical experience, a series of variables were selected to construct a prediction model based on the SVM machine learning algorithm.

Results: Among the 369 patients, preoperative DVT was observed in 21 patients (5.7%). The Multivariate regression analysis showed the following 5 independent factors associated with preoperative DVT: preoperative fibrinogen odds ratio [OR] = 7.306), age (OR = 1.133), history of hypertension (OR = 3.848), preoperative hematocrit (OR = 0.315), and D-dimer (OR = 2.032). The SVM model achieved a maximum and average area under the receiver operating characteristic curve (AUC) of 0.94 and 0.77 in the 10-fold cross-validation. Meanwhile, the accuracy, precision, and recall of the model were 0.98, 0.92, and 0.93, respectively. Additionally, the confusion matrix showed the classification results of the discriminant analysis.

Conclusions: SVM machine modeling is a promising method for the prediction of the risk of DVT in non-fractured patients awaiting TJA. However, future external validation is needed.

1. Background

Lower extremity total joint arthroplasty (TJA), including total knee arthroplasty (TKA) and total hip arthroplasty (THA), is an effective treatment for advanced joint disease[1, 2]. With surgical advancements in TJA, the incidence of complications after TJA has considerably decreased[3, 4]. Venous thromboembolism (VTE) has been a concern for orthopedists for a long period of time[5–7]. VTE includes deep vein thrombosis (DVT) and pulmonary embolism (PE), which are characterized by high incidence, high mortality, and high morbidity[8, 9]. In clinical practice, VTE can cause serious and potentially fatal complications and lead to fatal outcomes. Major orthopedic procedures include THA, TKA, and hip fracture surgery (HFS), which are regarded as important risk factors for postoperative VTE[10–12].

It has been previously shown that the incidence of postoperative VTE without prophylaxis can reach 42–57%[8], 41–85%[8], and 46–60%[13], respectively. In recent years, with the recommendation of guidelines and the emphasis of clinicians on thromboprophylaxis, the incidence of postoperative DVT in major

orthopedic surgery has shown a decreasing trend. Remarkably, DVT incidence following TKA ranged from about 18.1 to 48.6% without thromboprophylaxis in Asian patients[8].

At present, many studies have reported the risk factors for DVT after TJA, including patients' own factors (age, body mass index (BMI), diabetes, history of smoking, etc.), surgical factors (surgical approach, tourniquet, bone cement, anaesthesia modality, etc.) and perioperative management (pain control, prevention of postoperative venous thromboembolism, and early rehabilitation intervention)[14]. We note that, since the patient's own factors were already present before admission, these factors are important contributors to the hypercoagulable state of the patient's blood preoperatively. Therefore, some scholars speculated whether DVT had formed before joint arthroplasty, and it was verified by imaging screening. Wakabayashi et al[15] performed preoperative color ultrasound screening in 322 patients undergoing knee arthroplasty and confirmed the presence of DVT in 56 cases, the incidence of which was as high as 17.4% (56/322), including 3 cases with proximal thrombus. They found that comorbid rheumatoid arthritis, connective tissue disease, and planned revision knee arthroplasty were independent risk factors for DVT formation prior to TKA. Wakabayashi et al[16] performed preoperative vascular ultrasonography of the lower extremities in 505 patients undergoing THA, 62 of whom had DVT, which occurred in 12.3% (62/505). This research screened and identified related risk factors, including advanced age, history of major surgery, and undergoing revision THA. To date, adequate attention has been paid to the prevention and treatment of DVT following major orthopedic surgery, but there are few studies have reported the characteristics of preoperative DVT in non-fractured patients awaiting TJA in China. Meanwhile, there is a lack of a reliable prediction tool so far. A supervised machine learning method, the support vector machine (SVM), has demonstrated high performance in solving classification problems in many biomedical fields[17–19]. This technique has recently been used to develop an automated classification of diseases and to improve methods for detecting diseases in the clinical setting.

This study aims to investigate the incidence of preoperative DVT and develop a concise and reliable model for evaluation of the risk of preoperative DVT in patients awaiting total joint arthroplasty using the SVM method.

2. Methods

2.1 Patients

The Institutional Review Board (IRB) of the healthcare organization reviewed and approved the study. Because this was a retrospective analysis, the IRB confirmed the requirement for informed consent was waived. From March 2015 to August 2020, 431 patients undergoing elective TJA were obtained and analyzed. The inclusion and exclusion criteria are as outlined below. (1) Inclusion criteria: (a) Patients undergoing elective TJA; (b) Preoperative thrombus screening of bilateral lower extremity veins was performed using lower extremity vein color Doppler sonography; (c) Incidence of preoperative DVT, risk factors, and postoperative changes in thrombotic outcomes. (2) Exclusion criteria: (a) Presence of VTE was informed on admission; (b) Venous thromboembolism had recently occurred and was receiving

treatment; (c) Those with a fresh, old hip fracture, or fractures elsewhere in the lower extremity requiring primary or revision arthroplasty surgery; (d) Lower extremity tissue damage precludes venous ultrasonography from being done normally; (e) Incomplete clinical and imaging data.

2.2 Data collection

Clinical and radiographic data of patients who underwent elective TJA at our institution from March 2015 to August 2020 were collected. The clinical information was recorded by the patient's physician in charge before discharge and included age, sex, body mass index (BMI), diabetes, hypertension, malignancy, stroke, ischaemic heart disease, presence of lower limb varicose veins, congestive heart failure, preoperative diagnosis, surgical procedure, as well as laboratory tests, blood routine testing, coagulation function testing, liver function testing, renal function, and electrolyte examination. Here, D-dimer ≥ 0.5 *μg*/ml in joint surgery clinical practice was defined as *positive.*

2.3 Preoperative ultrasound examination of lower extremities

All ultrasound examinations performed on patients entailed the use of a Sonosite M Turbo (Sonosite, Bothell, WA, USA) (or similar device) with a 7–12 MHz 38 mm linear probe 1–3 days before surgery. All examinations were performed by the same sonographers. During the examination, the patient underwent mild abduction and external rotation of both lower extremities, starting from the midpoint of the inguinal ligament, and sequentially examined bilateral common femoral, femoral, popliteal, posterior tibial, peroneal, and intermuscular veins. In those with more difficult exposures, the popliteal, peroneal, and intermuscular veins may be further screened in the prone position. For cases with suspected thrombus, the diagnostic criteria are as follows[20]: I) the wall of the vein can not be compressed or only partially compressed, II) no blood flow signal or only partial blood flow signal is seen on color Doppler. Once thrombosis was confirmed, information on the anatomical location, side, presence or absence of symptoms of thrombosis was recorded.

2.4. Support vector machine

As a supervised machine learning algorithm, SVM has been widely applied many high-dimensional data analyses. Currently, SVM algorithms have shown satisfactory performance in medical data mining and bioinformatics. The algorithm performs classification by constructing a hyperplane in a high-dimensional space that differentiates the two classifications by finding a boundary between the two data clusters[21–23]. In this study, the SVM algorithm obtained a good classification performance by transforming the input space into a high-dimensional space using one nonlinear function called the kernel function. Figure 1 shows an example of an inseparable two-dimensional space that becomes separable after transforming the input space from low-dimensional to high-dimensional. To optimize training, we further split the training set into 80% (training) and 20% (development) stratified random sets. The input feature format can be recorded as "csv" or "xlxs" and imported to SVM. The training was performed using 10-fold cross-validation. Using an RBF kernel, the SVM needs to adjust two parameters, the error penalty parameter C and the γ coefficient, through grid search. The Linear kernel was chosen as the kernel

function in SVM. In addition, the C parameter indicates the degree of avoid misclassifying each training example during SVM optimization, where the value of C is set at 1.2. Data mining and analysis were performed using programs written in the Python programming language (Python 3.8.0, Python Software Foundation, https://www.python.org/). The area under the receiver operating characteristic curve (AUC), accuracy (ACC), precision, recall and confusion matrices were calculated to assess the performance of the SVM prediction model. These indicators are determined by the following equations:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

In these equations, TP: true positive; TN: true negative; FP: false positive; and FN: false negative. In the training process, tuning was considered for ML-based models to avoid overfitting and the best hyper-parameter for ML models was 10-fold cross-validation (CV). The code used to develop the SVM algorithm is shown the supplementary material.

2.5. Statistical analysis

Data analysis was performed using the SPSS 26.0 software package (SPSS, Chicago, IL). Continuous variables were presented as means ± standard deviations and categorical variables as frequency (percentage). If the distributions of continuous variables were normal, Student's t tests were applied. Conversely, if normality tests fail, Mann– Whitney tests are used. The $\chi 2$ tests or Fisher exact tests were used for comparison of categorical variables.

For clinical studies, it is necessary to understand the relationship between independent variables and dependent variables. In this study, univariate and multivariate analysis were used to identify the independent risk factors and calculate the odds ratio (OR) value. Multivariate logistic regression was performed to identify possible risk factors that are independently associated with the incidence of preoperative DVT, using a forward stepwise selection approach. Next, variables with an outcome with a $P < 0.05$ on univariate analysis were included in the multivariate logistic regression analysis. Taking into account the number of available events, it is necessary to keep the simplicity of the final regression model in the process of setting variables. To facilitate the application of machine learning models, prevent dimensional catastrophe and reduce training time, subsets of relevant features are used for feature selection. At the same time, it is important to note in this study that the features selected by the machine learning model need not necessarily be exactly the same as in multivariate regression logistic regression.

3. Results

The baseline characteristics of the patients are demonstrated in Table 1. Among the 369 patients, we observed 21 patients (5.7%) with preoperative DVT (DVT group) and 348 (94.3%) patients without DVT (Non-DVT group). The mean age of total patients was 62.6 ± 7.4 years, and 276 patients (74.8%) were female. Between the DVT group and the Non-DVT group, there were no statistically significant differences, including sex, body mass index (BMI), diabetes, stroke, malignancy, presence of lower limb varicose veins, preoperative diagnosis, surgical site, preoperative white blood cell (WBC), preoperative prothrombin time (PT), preoperative activated partial thromboplastin time (APTT), preoperative albumin (ALB), preoperative triacylglycerol (TG), preoperative total cholesterol (TC), preoperative low-density lipoprotein (LDL-C), preoperative apolipoprotein A (apoA), and preoperative apolipoprotein B (apoB). In this study, we found the mean age in the DVT [mean, SD. 67.1 ± 10.4] group was significantly higher than that in the Non-DVT group [mean, SD. 62.3 ± 7.1]. The preoperative hemoglobin (Hb), hematocrit (Hct), thrombin time (TT), and international normalized ratio (INR) were lower in the DVT group than in the Non-DVT group. We found a significantly higher proportion of history of hypertension results in the Non-DVT group (16/76.2%) than in the Non-DVT group. The preoperative fibrinogen in the DVT group [mean, SD. 3.7 (1.1)] was higher compared with that in the Non-DVT group [mean, SD. 2.9 (0.4)]. Meanwhile, the preoperative high-density lipoproteins (HDL-C) was lower in the Non-DVT group [mean, SD. 1.1 (0.2)] than in the DVT group [mean, SD. 1.3 (0.3)] ($p < 0.001$). Notably, we noticed that a significantly higher proportion of D-dimer positivity results in the DVT group (16/76.2%) than in the Non-DVT group (167/48.4%) ($p = 0.022$).

Table 1
Baseline characteristics of patients.

		Total	Non-DVT group	DVT group	P value
Number of patients		369	348	21	
Age (year)		62.6 (7.4)	62.3 (7.1)	67.1 (10.4)	0.003
Sex (%)					
	Fmale	276 (74.8)	258 (74.1)	18 (85.7)	0.354
	Male	93 (25.2)	90 (25.9)	3 (14.3)	
BMI (kg/m^2)		25.7 (4.1)	25.7 (4.1)	25.3 (3.9)	0.646
History of hypertension (%)					
	No	195 (52.8)	190 (54.6)	5 (23.8)	0.012
	Yes	174 (47.2)	158 (45.4)	16 (76.2)	
Diabetes (%)					
	No	299 (81.0)	281 (80.7)	18 (85.7)	0.782
	Yes	70 (19.0)	67 (19.3)	3 (14.3)	
Stroke (%)					
	No	361 (97.8)	341 (98.0)	20 (95.2)	0.945
	Yes	8 (2.2)	7 (2.0)	1 (4.8)	
History of malignancy (%)					
	No	360 (97.6)	339 (97.4)	21 (100.0)	0.986
	Yes	9 (2.4)	9 (2.6)	0 (0.0)	

Abbreviations: DVT: deep vein thrombosis; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OA: osteoarthritis; FHN: femoral head necrosis; RA: rheumatoid arthritis; AS: ankylosing spondylitis; AD: acetabular dysplasia; TA: traumatic arthritis.

		Total	Non-DVT group	DVT group	P value
Presence of lower limb varicose veins (%)					
	No	354 (95.9)	334 (96.0)	20 (95.2)	1
	Yes	15 (4.1)	14 (4.0)	1 (4.8)	
Preoperative diagnosis (%)					
	OA	197 (53.4)	185 (53.2)	12 (57.1)	0.74
	FHN	99 (26.8)	93 (26.7)	6 (28.6)	
	RA & AS	24 (6.5)	24 (6.9)	0 (0.0)	
	AD	25 (6.8)	24 (6.9)	1 (4.8)	
	TA	24 (6.5)	22 (6.3)	2 (9.5)	
Surgical procedure (%)					
	Primary Total Joint Arthroplasty	349 (94.6)	331 (95.1)	18 (85.7)	0.177
	Revision Total Joint Arthroplasty	20 (5.4)	17 (4.9)	3 (14.3)	
Surgical site (%)					
	Hip	158 (42.8)	152 (43.7)	6 (28.6)	0.258
	knee	211 (57.2)	196 (56.3)	15 (71.4)	
Preoperative Hct (%)		37.0 (1.6)	37.0 (1.6)	36.0 (1.6)	0.005
Preoperative Hb (g/L)		148.1 (9.2)	150.0 (4.8)	116.4 (6.9)	< 0.001

Abbreviations: DVT: deep vein thrombosis; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OA: osteoarthritis; FHN: femoral head necrosis; RA: rheumatoid arthritis; AS: ankylosing spondylitis; AD: acetabular dysplasia; TA: traumatic arthritis.

		Total	Non-DVT group	DVT group	P value
Preoperative WBC (10*^{9}/L)		7.3 (0.5)	7.3 (0.5)	7.2 (0.3)	0.344
Preoperative PT (s)		11.9 (1.1)	11.9 (1.1)	11.6 (0.6)	0.254
Preoperative APTT (s)		28.8 (6.3)	28.9 (6.4)	28.1 (3.0)	0.599
Preoperative TT (s)		18.1 (3.0)	18.2 (3.0)	16.3 (1.3)	0.004
Preoperative INR		1.13 (0.1)	1.2 (0.1)	1.08 (0.1)	0.141
Preoperative FIB (mg/dL)		2.9 (0.5)	2.9 (0.4)	3.7 (1.1)	< 0.001
Preoperative D-dimer (%)					
	Negative	186 (50.4)	181 (52.0)	5 (23.8)	0.022
	Positive	183 (49.6)	167 (48.0)	16 (76.2)	
Preoperative ALB (g/L)		40.1 (3.6)	40.1 (3.6)	40.0 (3.2)	0.977
Preoperative TG (mmol/L)		1.5 (0.6)	1.5 (0.6)	1.8 (1.0)	0.057
Preoperative TC (mmol/L)		4.5 (0.9)	4.5 (0.9)	4.4 (1.0)	0.929
Preoperative HDL-C (mmol/L)		1.1 (0.2)	1.1 (0.2)	1.2 (0.3)	0.761
Preoperative LDL-C (mmol/L)		2.5 (0.5)	2.5 (0.5)	2.3 (0.7)	0.077
Preoperative apoA (mmol/L)		1.3 (0.1)	1.3 (0.1)	1.3 (0.3)	0.096
Preoperative apoB (mmol/L)		1.0 (0.1)	1.0 (0.1)	1.0 (0.1)	0.55

Abbreviations: DVT: deep vein thrombosis; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OA: osteoarthritis; FHN: femoral head necrosis; RA: rheumatoid arthritis; AS: ankylosing spondylitis; AD: acetabular dysplasia; TA: traumatic arthritis.

The univariate and multivariate logistic regression analyses (Table 2) showed that age, history of hypertension, preoperative Hct, preoperative fibrinogen, and preoperative D-dimer were independently associated with DVT. Next, these independent parameters were imported into the SVM model. To concisely summarize the prediction performance of the SVM model, we constructed ROC curves, which evaluate the performance of a model in a way that takes the uncertainty of each prediction into account. In parallel, we calculated the ACC, precision, and recall of the SVM model. Based on the overall dataset, the ROC curve of the prediction model is shown in Figure. 2a, and the AUC was 0.94. Concurrently, the precision, ACC, and recall of the prediction model were 0.98, 0.92, and 0.93, respectively. The confusion matrix shows the classification results of the discriminant analysis (Figure. 2b). In this paper, the prediction performance of the SVM model was verified by 10-fold cross-validation (CV). The max AUC was 0.94 and the average AUC of the 10-fold CV was 0.77. Figure 3 shows the AUC value of the 10-fold CV in the training process. In this research, the relative importance of variables in the SVM model was shown in Fig. 4. The importance of the variables in the SVM model is in decreasing order as follows: preoperative fibrinogen, age, preoperative D-dimer, history of hypertension, and preoperative Hct.

Table 2
Univariate and multivariate logistic regression model analyses of preoperative DVT in non-fractured patients awaiting TJA.

	Univariate analysis		**Multivariate analysis**	
	OR (95% CI)	P value	OR (95% CI)	P value
Age (year)	1.092 (1.028–1.159)	0.004	1.133 (1.044–1.231)	0.003
Sex (%)	0.478 (0.137–1.660)	0.245	NA	
BMI (kg/m^2)	0.975 (0.874–1.087)	0.645	NA	
History of hypertension (%)	3.848 (1.379–10.737)	0.01	4.152 (1.176–14.660)	0.027
Diabetes (%)	0.699 (0.200-2.442)	0.575	NA	
Stroke (%)	2.436 (0.286–20.770)	0.416	NA	
History of malignancy (%)	0.000 (0.000-Inf)	0.991	NA	
Presence of lower limb varicose veins (%)	1.193 (0.149–9.532)	0.868	NA	
Preoperative diagnosis (%)	0.961 (0.658–1.404)	0.837	NA	
Surgical procedure (%)	3.245 (0.870-12.097)	0.08	NA	
Surgical site (%)	1.939 (0.735–5.115)	0.181	NA	
Preoperative Hct (%)	0.670 (0.503–0.891)	0.006	0.315 (0.201–0.498)	0.031
Preoperative Hb (g/L)	0.000 (0.000-Inf)	1	NA	
Preoperative WBC ($10*^9/L$)	0.660 (0.279–1.558)	0.343	NA	
Preoperative PT (s)	0.794 (0.534–1.180)	0.254	NA	
Preoperative APTT (s)	0.981 (0.913–1.054)	0.598	NA	
Preoperative TT (s)	0.807 (0.695–0.938)	0.005	NA	
Preoperative INR	69.632 (0.244-19910.466)	0.141	NA	

Abbreviations: DVT: deep vein thrombosis; TJA: total joint arthroplasty; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OR: odds ratio; CI, confidence interval; NA: not available.

	Univariate analysis		Multivariate analysis	
Preoperative FIB (mg/dL)	13.467 (5.094–35.599)	< 0.001	7.306 (2.653 – 0.115)	< 0.001
Preoperative D dimer (%)	3.468 (1.243–9.676)	0.018	2.032(1.003–6.031)	0.043
Preoperative ALB (g/L)	0.998 (0.884–1.127)	0.977	NA	
Preoperative TG (mmol/L)	1.937 (0.979–3.832)	0.058	NA	
Preoperative TC (mmol/L)	0.977 (0.585–1.633)	0.929	NA	
Preoperative HDL-C (mmol/L)	22.602 (1.690-302.305)	0.089	NA	
Preoperative LDL-C (mmol/L)	0.475 (0.208–1.087)	0.078	NA	
Preoperative apoA (mmol/L)	21.117 (0.591-754.634)	0.095	NA	
Preoperative apoB (mmol/L)	2.606 (0.113–59.873)	0.549	NA	

Abbreviations: DVT: deep vein thrombosis; TJA: total joint arthroplasty; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OR: odds ratio; CI, confidence interval; NA: not available.

4. Discussion

In this work, we investigated the relationship between clinical variables and DVT, and identified the independent risk factors using multivariate logistic regression involving 369 consecutive patients. Subsequently, we developed and validated an SVM algorithm to predict the risk of preoperative DVT in non-fractured patients awaiting TJA. A range of model evaluation indexes indicated that the SVM model could deliver satisfactory performance and had good clinical application value and promotion value[17, 18, 24].

Although the incidence of and potential risk factors for preoperative DVT in arthroplasty patients have been documented previously, the diagnoses in this population have all been fractures, particularly those with a predominance of hip fractures. Song et al.[25] analyzed the results of preoperative lower extremity deep venography in 119 patients with femoral neck fractures, in whom thrombi were found in 35 cases, for an incidence rate of 29.4%. A similar study was subsequently conducted by Xia et al.[26], which found that thrombotic and pulmonary embolic incidences of 18.9% and 1%, respectively, were found in the preoperative evaluation of 301 patients with femoral neck fractures. Luksameearunothai et al[27] examined 92 hip fracture patients with preoperative imaging studies and found that the incidence of DVT was 16.3%. Of interest is the presence of several high-risk factors associated with DVT, such as fracture,

advanced age, and continuous ambulation before elective arthroplasty in most hip fracture populations. Consequently, we cautiously assume that DVT may be present preoperatively in many patients. However, we know little about this.

The incidence of preoperative DVT in the non-fractured population undergoing elective joint arthroplasty in this study was 5.7%, which is in agreement with previous studies[27–29]. In a separate study, Kim et al. [29] investigated 311 osteoarthritis patients for DVT before TKA and found that the incidence of preoperative DVT was 4.5%. In addition, it has also been documented that the incidence of preoperative thrombus was higher than in this study. Watanabe et al[29] used computed tomography (CT) in 71 patients undergoing TKA to screen for preoperative and postoperative thrombophilia, which showed an 8.0 % incidence of preoperative thromboembolism. The reason why the literature reported that thrombosis occurred differently in the preoperative non-fractured population may be related to the demographic characteristics of the study population as well as differences in medical history. Until now, there has been some additional research on the prediction of DVT. Frustratingly, it is difficult to accurately predict the incidence of preoperative DVT and identify related risk factors[28, 30]. In this report, we developed and validated an SVM model for preoperative DVT in non-fractured patients awaiting TJA based on a machine learning algorithm.

In the present study, we identified five factors that were independently associated with preoperative DVT. Fibrinogen was an independent predictor. Fibrinogen is an inflammatory protein that gets converted to fibrin in the presence of thrombin and directly influences platelet adhesion and activation. Meanwhile, fibrinogen is a soluble plasma protein that plays an important function during clot formation. In a previous retrospective study, fibrinogen was a significant risk factor for preoperative DVT in patients with lower extremity fractures, playing an important role in the preoperative evaluation of patients with lower extremity fractures[31, 32]. Similar conclusions were also obtained in the present study. Multivariate logistic regression analysis revealed fibrinogen as an independent risk factor (OR = 7.306, 95%CI = 2.653–20.115, $P < 0.001$). This has important implications for our preoperative preparation and the relatively quick and concise identification of high-risk individuals.

Previous literature suggested that age was an important risk factor for DVT. The incidence of DVT also gradually increased with age, from 5 to 89 years, and the incidence of DVT increased by 5‰ – 6‰ per additional year[33]. Advanced age was observed to be a risk factor for preoperative DVT in non-fractured patients awaiting TJA in our study. Probably, patients of advanced age generally have a larger proportion of the underlying disease, which is prone to pathological changes such as vascular endothelial injury. However, the specific mechanism needs further in-depth study.

The results of this study found that a positive D-dimer and a history of hypertension were risk factors for preoperative DVT in non-fractured patients awaiting TJA. In clinical practice, elevated D-dimer levels are sensitive for the detection of VTE, but lack specificity, and elevated D-dimer levels have been observed in conditions such as trauma, inflammation, infection, and tumors. Therefore, although D-dimer elevation is somewhat helpful for initial screening for DVT, the sample size needs to be enlarged in subsequent

studies to seek an appropriate D-dimer cut-off value to assist in the exclusion of thrombus. A prospective, multicenter investigation with a large sample is essential in the future. Previous literature has confirmed that a history of hypertension can increase the risk of VTE after orthopedic surgery[34–36]. A meta-analysis of 16 studies involving orthopedic surgical patients (68955 males and 53057 females) reported a significant association between hypertension and postoperative DVT (OR = 2.89, 95%CI = 2.18–3.83, P < 0.05, Z = 7.38)[37]. The results of this study suggest that the prevalence of DVT is also significantly higher in patients with hypertension prior to joint arthroplasty, and the presumed cause is related to the fact that patients with hypertension are more likely to have disorders of the coagulation system, vascular inflammation, and endothelial dysfunction. One highlight of our work was using the SVM machine learning technique to predict the risk of preoperative DVT in non-fractured patients awaiting TJA from routinely available variables. The 10-fold CV AUC, ACC, precision, recall and confusion matrix indicated that our model performed well in this research. Thus, the SVM model could be used as a reliable tool for distinguishing non-fractured patients awaiting TJA at high risk of preoperative DVT and may provide useful information for clinicians to optimize individual therapy management.

However, the limitations of this study should be stated. First, the nature of a retrospective study might have resulted in selection bias. Second, to be accurate and effective, the SVM algorithm should be trained on a high quantity of data, which needs to be further validated in other regions and medical centers. The ML algorithm model we established, to some extent, was confined to one single institution, which might restrict its generalizability pending further validation in real-world scenarios. Third, optimization and validation of the model are based on artificial intelligence techniques, which present new challenges for hospitals and clinicians. Finally, although we have made every attempt to collect and analyze as much clinical data as possible to identify the risk factors for preoperative DVT, some variables could not be explored because of missing data.

5. Conclusions

In confusion, we developed and validated an SVM model for individualized prediction of the risk of preoperative DVT in non-fractured patients awaiting TJA by utilizing readily available preoperative variables. The SVM model demonstrated satisfactory performance in this study. Additionally, the SVM prediction model can accurately identify whether patients are at high-risk of preoperative DVT and can provide an easy tool for the treatment team to make precise decisions. In the future, we hope to continue to optimize the model and integrate imaging and molecular data to further improve the performance of our model and better support clinical decision-making.

Abbreviations

DVT: deep vein thrombosis; BMI, body mass index; Hct: hematocrit; Hb: hemoglobin; WBC: white blood cell; ALB: albumin; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; INR: international normalized ratio; TG: triacylglycerol; TC: total cholesterol; HDL-C: low density lipoprotein; HDL-C: high-density lipoproteins; apoA: apolipoprotein A; apoB: apolipoprotein B; OA:

osteoarthritis; FHN: femoral head necrosis; RA: rheumatoid arthritis; AS: ankylosing spondylitis; AD: acetabular dysplasia; TA: traumatic arthritis, SVM: support vector machine; TKA: total knee arthroplasty; THA: total hip arthroplasty; VTE: venous thromboembolism; PE: pulmonary embolism; HFS: hip fracture surgery; RBF: radial basis function; ACC: accuracy; AUC: area under the receiver operating characteristic curve; CV: cross-validation; TP: true positive; TN: true negative; FP: false positive; FN: false negative.

Declarations

Ethics approval and consent to participate

The study was approved by an institutional ethics committee at the Taizhou Central Hospital (Affiliated Hospital to Taizhou College). Considering that this work was a retrospective study, the ethics committee waived the requirement for informed consent from patients.

Consent for publication

Not applicable.

Availability of data and materials

The data set supporting the conclusion of this article is available on request to the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

No funds were received in support of this work.

Authors' contributions

HSW collected the data, analyzed the data, and drafted the manuscript. MYY supervised the project and reviewed the manuscript. FJZ, WLL, TTF, and QML conceived of the study, participated in its design and coordination, and helped to draft the manuscript. MYY was responsible for the whole project, designed the study, and supervised the study. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Bannister G, Young S, Baker A, Mackinnon J, Magnusson P: **Control of bleeding in cemented arthroplasty**. *The Journal of bone and joint surgery British volume* 1990, **72**(3):444-446.

2. Verstraete M: **Clinical application of inhibitors of fibrinolysis**. *Drugs* 1985, **29**(3):236-261.
3. Austin MS, Parvizi J, Grossman S, Restrepo C, Klein GR, Rothman RH: **The inferior vena cava filter is effective in preventing fatal pulmonary embolus after hip and knee arthroplasties**. *The Journal of arthroplasty* 2007, **22**(3):343-348.
4. Mismetti P, Laporte S, Zufferey P, Epinat M, Decousus H, Cucherat M: **Prevention of venous thromboembolism in orthopedic surgery with vitamin K antagonists: a meta-analysis**. *Journal of Thrombosis and Haemostasis* 2004, **2**(7):1058-1070.
5. Bagaria V, Modi N, Panghate A, Vaidya S: **Incidence and risk factors for development of venous thromboembolism in Indian patients undergoing major orthopaedic surgery: results of a prospective study**. *Postgraduate medical journal* 2006, **82**(964):136-139.
6. Piovella F, Wang CJ, Lu H, Lee K, Lee L, Lee W, Turpie A, Gallus A, Planès A, Passera R: **Deep-vein thrombosis rates after major orthopedic surgery in Asia. An epidemiological study based on postoperative screening with centrally adjudicated bilateral venography**. *Journal of Thrombosis and Haemostasis* 2005, **3**(12):2664-2670.
7. Xu X, Jiang J, Liu W, Li X, Lu H: **Application of thromboelastography to evaluate the effect of different routes administration of tranexamic acid on coagulation function in total hip arthroplasty**. *J Orthop Surg Res* 2019, **14**(1):430-430.
8. Stein PD, Kayali F, Olson RE, Milford CE: **Pulmonary thromboembolism in Asians/Pacific islanders in the United States: analysis of data from the National Hospital Discharge Survey and the United States Bureau of the Census**. *The American journal of medicine* 2004, **116**(7):435-442.
9. Won YS, Kim M, Jun KW, Nam WS, Ahn S, Hwang J-K, Kim S-D, Park S-C, Yun SS, Lee W-C: **Incidence and Clinical Characteristics of Deep Vein Thrombosis (DVT) after Total Knee Arthroplasty (TKA) with DVT Chemoprophylaxis**. *World Journal of Cardiovascular Diseases* 2014, **4**(11):531.
10. Ahmed O, Kim YJ, Patel MV, Luu HH, Scott B, Cohen K: **Efficacy and Safety of Mechanical IVC Filtration for Preventing Pulmonary Embolism in High-Risk Orthopedic Patients Undergoing Total Hip or Knee Arthroplasty**. *The Journal of Arthroplasty* 2021.
11. Dhillon K, Askander A, Doraisamy S: **Postoperative deep-vein thrombosis in Asian patients is not a rarity: a prospective study of 88 patients with no prophylaxis**. *The Journal of bone and joint surgery British volume* 1996, **78**(3):427-430.
12. Leizorovicz A, Turpie A, Cohen A, Wong L, Yoo M, Dans A, Group SS: **Epidemiology of venous thromboembolism in Asian patients undergoing major orthopedic surgery without thromboprophylaxis. The SMART study**. *Journal of Thrombosis and Haemostasis* 2005, **3**(1):28-34.
13. Geerts WH, Pineo GF, Heit JA, Bergqvist D, Lassen MR, Colwell CW, Ray JG: **Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy**. *Chest* 2004, **126**(3):338S-400S.
14. Geerts WH, Heit JA, Clagett GP, Pineo GF, Colwell CW, Anderson FA, Wheeler HB: **Prevention of venous thromboembolism**. *Chest* 2001, **119**(1):132S-175S.

15. Wakabayashi H, Hasegawa M, Niimi R, Yamaguchi T, Naito Y, Sudo A: **The risk factor of preoperative deep vein thrombosis in patients undergoing total knee arthroplasty**. *Journal of Orthopaedic Science* 2017, **22**(4):698-702.
16. Wakabayashi H, Hasegawa M, Niimi R, Sudo A: **Clinical analysis of preoperative deep vein thrombosis risk factors in patients undergoing total hip arthroplasty**. *Thrombosis research* 2015, **136**(5):855-858.
17. Donaldson I, Martin J, De Bruijn B, Wolting C, Lay V, Tuekam B, Zhang S, Baskin B, Bader GD, Michalickova K: **PreBIND and Textomy–mining the biomedical literature for protein-protein interactions using a support vector machine**. *BMC bioinformatics* 2003, **4**(1):1-13.
18. Kim W, Kim KS, Lee JE, Noh D-Y, Kim S-W, Jung YS, Park MY, Park RW: **Development of novel breast cancer recurrence prediction model using support vector machine**. *Journal of breast cancer* 2012, **15**(2):230-238.
19. Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ: **Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes**. *BMC medical informatics and decision making* 2010, **10**(1):1-7.
20. Needleman L, Cronan JJ, Lilly MP, Merli GJ, Adhikari S, Hertzberg BS, DeJong MR, Streiff MB, Meissner MH: **Ultrasound for lower extremity deep venous thrombosis: multidisciplinary recommendations from the Society of Radiologists in Ultrasound Consensus Conference**. *Circulation* 2018, **137**(14):1505-1515.
21. Tang Z, Chen Y, Ye S, Hu R, Wang H, He J, Huang Q, Chang S: **Fully memristive spiking-neuron learning framework and its applications on pattern recognition and edge detection**. *Neurocomputing* 2020, **403**:80-87.
22. Tang Z, Zhu R, Lin P, He J, Wang H, Huang Q, Chang S, Ma Q: **A hardware friendly unsupervised memristive neural network with weight sharing mechanism**. *Neurocomputing* 2019, **332**:193-202.
23. Wu EQ, Lin C-T, Zhu L-M, Tang Z, Jie Y-W, Zhou G-R: **Fatigue Detection of Pilots' Brain Through Brains Cognitive Map and Multilayer Latent Incremental Learning Model**. *IEEE Transactions on Cybernetics* 2021.
24. Wang H, Tang Z-R, Li W, Fan T, Zhao J, Kang M, Dong R, Qu Y: **Prediction of the risk of C5 palsy after posterior laminectomy and fusion with cervical myelopathy using a support vector machine: an analysis of 184 consecutive patients**. *J Orthop Surg Res* 2021, **16**(1):1-9.
25. Song K, Yao Y, Rong Z, Shen Y, Zheng M, Jiang Q: **The preoperative incidence of deep vein thrombosis (DVT) and its correlation with postoperative DVT in patients undergoing elective surgery for femoral neck fractures**. *Archives of orthopaedic and trauma surgery* 2016, **136**(10):1459-1464.
26. Xia Z-N, Xiao K, Zhu W, Feng B, Zhang B-Z, Lin J, Qian W-W, Jin J, Gao N, Qiu G-X: **Risk assessment and management of preoperative venous thromboembolism following femoral neck fracture**. *J Orthop Surg Res* 2018, **13**(1):1-9.
27. Luksameearunothai K, Sa-Ngasoongsong P, Kulachote N, Thamyongkit S, Fuangfa P, Chanplakorn P, Woratanarat P, Suphachatwong C: **Usefulness of clinical predictors for preoperative screening of**

deep vein thrombosis in hip fractures. *BMC musculoskeletal disorders* 2017, **18**(1):1-7.

28. Kanchanabat B, Stapanavatr W, Meknavin S, Soorapanth C, Sumanasrethakul C, Kanchanasuttirak P: **Systematic review and meta-analysis on the rate of postoperative venous thromboembolism in orthopaedic surgery in Asian patients without thromboprophylaxis**. *Journal of British Surgery* 2011, **98**(10):1356-1364.
29. Kim K-I, Cho K-Y, Jin W, Khurana SS, Bae D-K: **Recent Korean perspective of deep vein thrombosis after total knee arthroplasty**. *The Journal of arthroplasty* 2011, **26**(7):1112-1116.
30. Chung L-H, Chen W-M, Chen C-F, Chen T-H, Liu C-L: **Deep vein thrombosis after total knee arthroplasty in asian patients without prophylactic anticoagulation**. *Orthopedics* 2011, **34**(1).
31. Lin C, Chen Y, Chen B, Zheng K, Luo X, Lin F: **D-Dimer Combined with Fibrinogen Predicts the Risk of Venous Thrombosis in Fracture Patients**. *Emergency Medicine International* 2020, **2020**.
32. Wen H, Chen Y: **The predictive value of platelet to lymphocyte ratio and D-dimer to fibrinogen ratio combined with WELLS score on lower extremity deep vein thrombosis in young patients with cerebral hemorrhage**. *Neurological Sciences* 2021:1-7.
33. Valerio L, Fedeli U, Schievano E, Avossa F, Barco S: **Decline in overall pulmonary embolism-related mortality and increasing prevalence of cancer-associated events in the Veneto region (Italy), 2008-2019**. *Thrombosis and Haemostasis* 2021(AAM).
34. Goel R, Patel EU, Cushing MM, Frank SM, Ness PM, Takemoto CM, Vasovic LV, Sheth S, Nellis ME, Shaz B: **Association of perioperative red blood cell transfusions with venous thromboembolism in a North American registry**. *JAMA surgery* 2018, **153**(9):826-833.
35. Kim K, Yamashita Y, Morimoto T, Kitai T, Yamane T, Ehara N, Kinoshita M, Kaji S, Amano H, Takase T: **Risk factors for major bleeding during prolonged anticoagulation therapy in patients with venous thromboembolism: from the COMMAND VTE Registry**. *Thrombosis and haemostasis* 2019, **119**(09):1498-1507.
36. Silvey M, Brandão LR: **Risk factors, prophylaxis, and treatment of venous thromboembolism in congenital heart disease patients**. *Frontiers in pediatrics* 2017, **5**:146.
37. Yang Y, Li Z, Liang H, Tian J: **Association between metabolic syndrome and venous thromboembolism after total joint arthroplasty: a meta-analysis of cohort studies**. *J Orthop Surg Res* 2020, **15**(1):1-10.

Figures

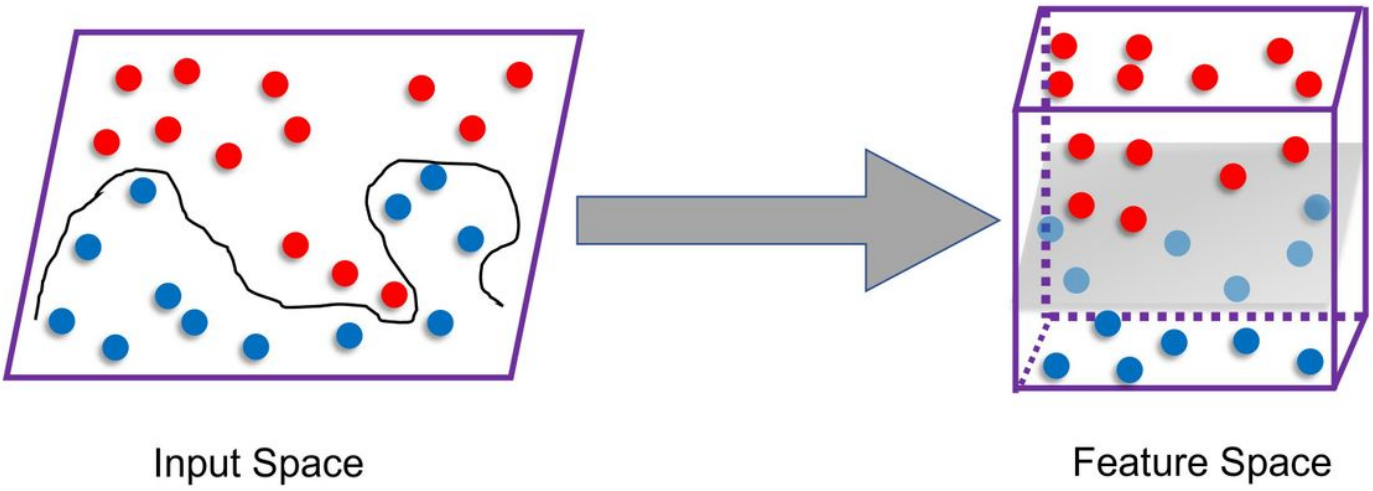

Figure 1

The illustration shows an example of an inseparable two-dimensional space that becomes separable after transforming the input space from low-dimensional to high-dimensional.

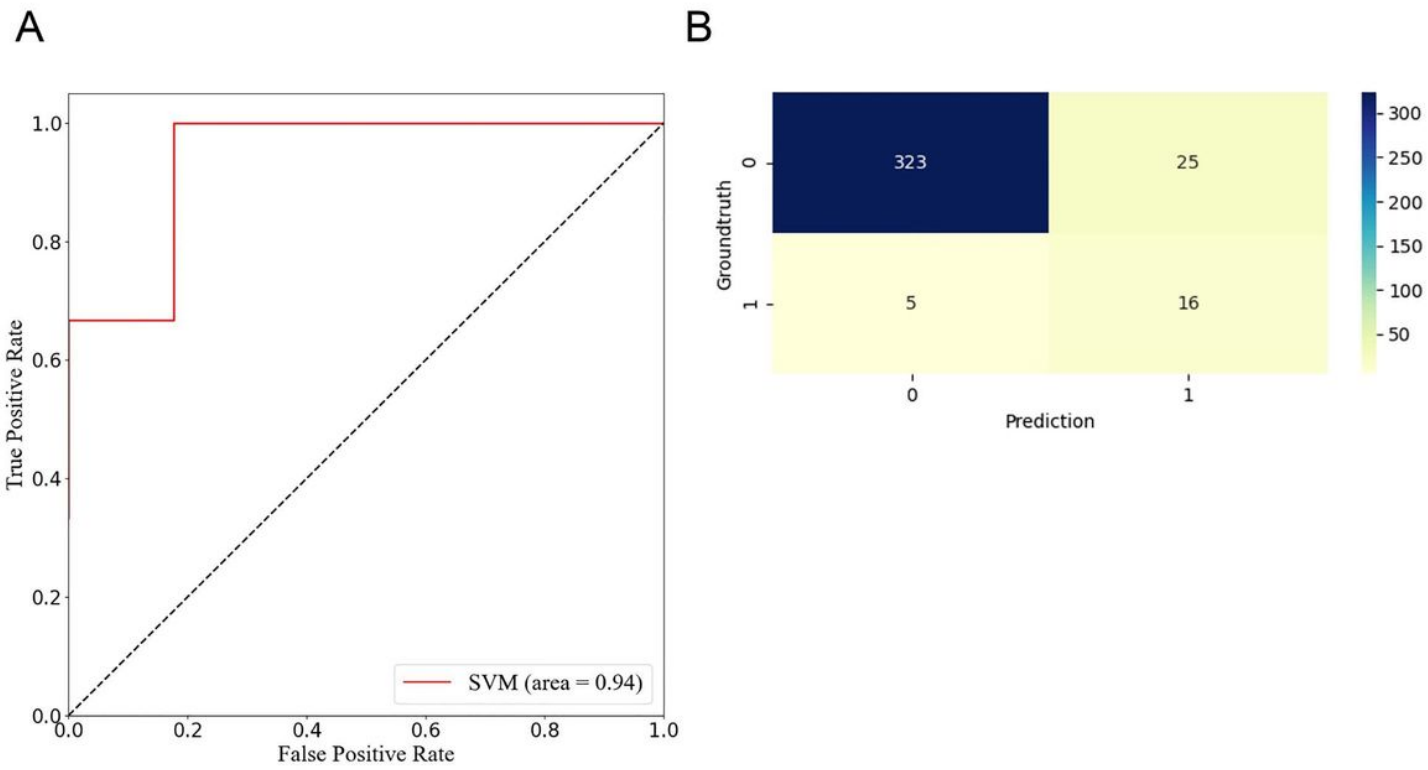

Figure 2

a. Receiver operating characteristic curve analysis of the SVM model; b. The results of the confusion matrix show good predictive ability of the SVM model.

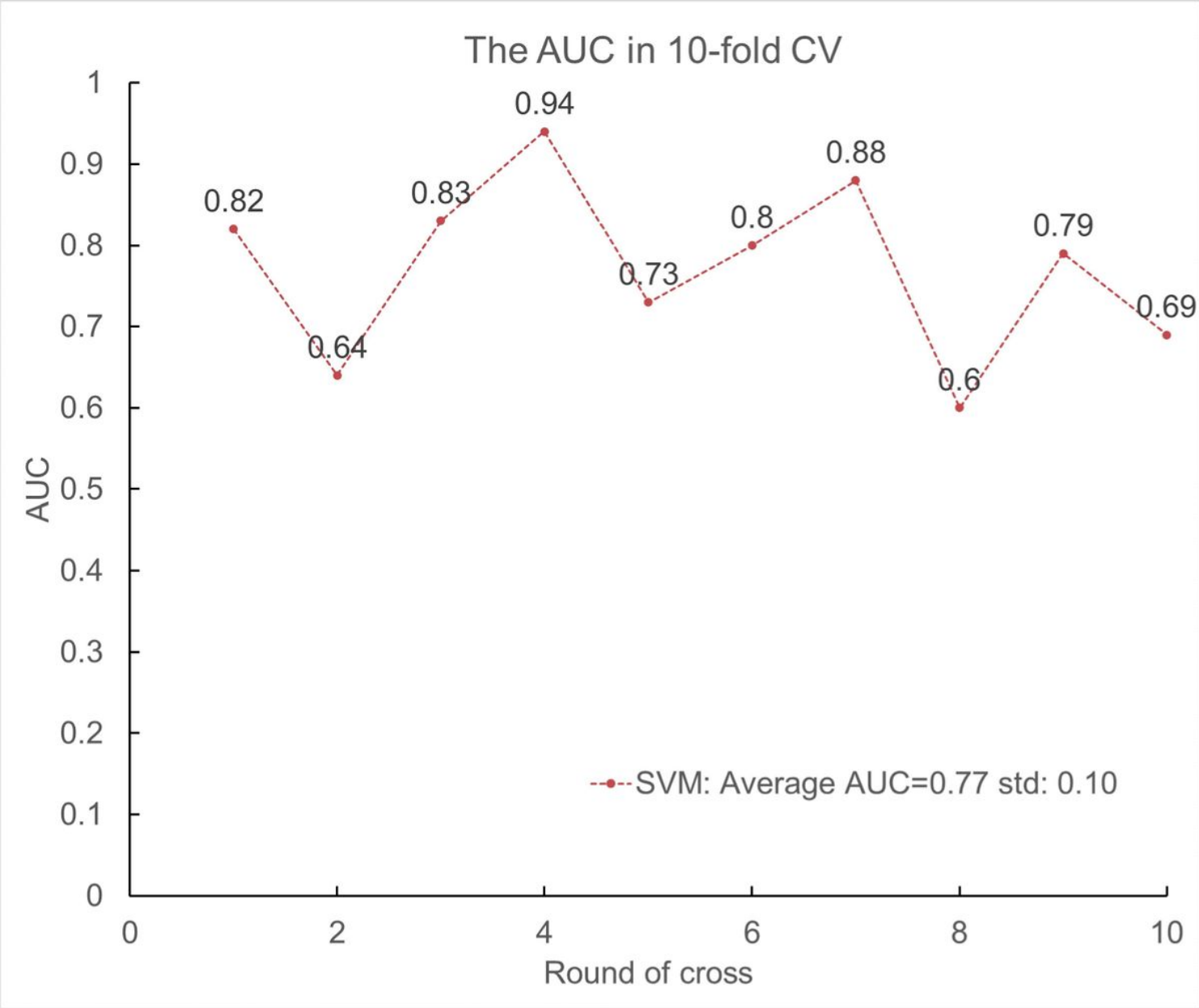

Figure 3

ROC curve analysis of a 10-fold CV of a support vector machine (SVM) algorithm to validate the performance of a model predicting the risk of preoperative DVT in non-fractured patients awaiting total joint arthroplasty.

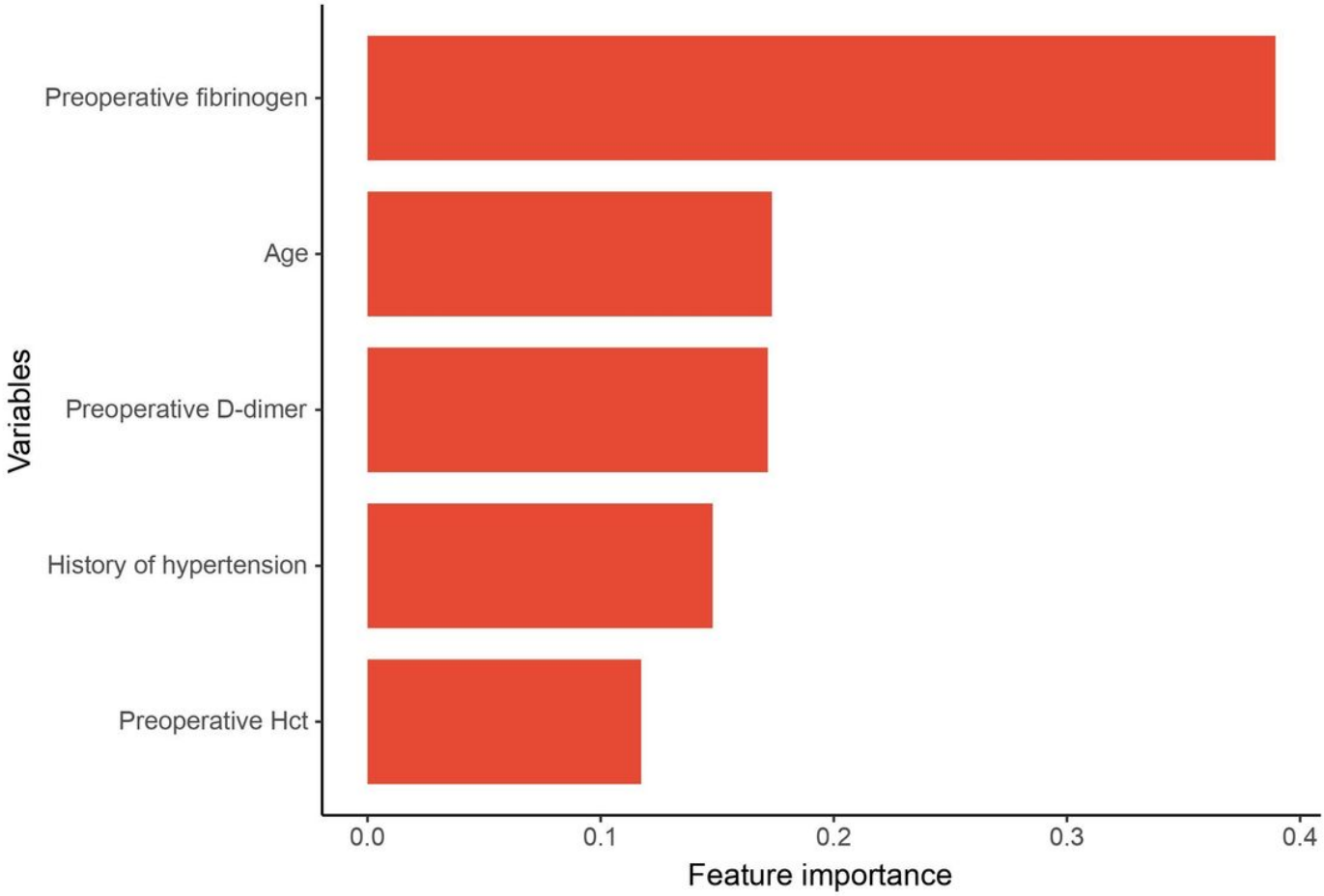

Figure 4

The importance of the variables in the SVM model is in decreasing order as follows: preoperative fibrinogen, age, preoperative D-dimer, history of hypertension, and preoperative Hct.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx