Predicting lower extremity deep vein thrombosis in elderly patients with hip fracture: A machine learning model with emphasis on diabetes as a key risk factor

QIAN SONG[1*], PING'AN SHI[2*], PENG TIAN[3*], AIJUN CHAO[1], YINGUANG ZHANG[3] and QIANG DONG[3]

[1]Department of Osteo-Internal Medicine, Tianjin University Tianjin Hospital, Tianjin 300122, P.R. China;
[2]Department of Endocrinology, Tianjin University Tianjin Hospital, Tianjin 300122, P.R. China;
[3]Department of Orthopedics, Tianjin University Tianjin Hospital, Tianjin 300122, P.R. China

Received February 25, 2026; Accepted June 15, 2026

DOI: 10.3892/etm.2026.13253

Abstract. Lower extremity deep vein thrombosis (DVT) is a serious complication in elderly patients with hip fracture, contributing to increased morbidity and mortality. Diabetes mellitus, with its prothrombotic state, may further elevate this risk. Early identification of high-risk patients is important for targeted thromboprophylaxis. The objective of the present study was to develop and validate machine learning (ML) models for predicting DVT using clinical variables available at admission in elderly patients with hip fracture, with a specific focus on diabetes as a key predictor. A retrospective cohort study of elderly patients (≥65 years) with hip fracture who were admitted to a tertiary academic medical center (Tianjin Hospital, China) between January 2020 and December 2025, was conducted. A total of seven ML algorithms were developed and validated using a 70-30 split with 10-fold cross-validation. Model interpretability was enhanced using Shapley Additive exPlanations (SHAP) analysis. Subgroup analyses were performed to evaluate model performance across diabetic and non-diabetic populations. DVT occurred in 16.5% (66/400) of patients during hospitalization. Diabetes was present in 32.5% of the cohort and was significantly associated with DVT (odds ratio=2.64; 95% CI: 1.56-4.48). The random forest model demonstrated an improved performance [area under the curve (AUC)=0.92; accuracy=0.87; sensitivity=0.85; specificity=0.88]. SHAP analysis identified diabetes-associated variables (hemoglobin A1c, diabetes duration and fasting glucose) among the top predictors, along with age, albumin, D-dimer and immobility on admission. Lipid parameters, including low-density lipoprotein cholesterol and triglycerides, also contributed to prediction. Model performance remained robust in diabetic (AUC=0.94) and non-diabetic (AUC=0.90) subgroups. Calibration was good (slope=1.02; intercept=-0.02; Brier score=0.09; Hosmer-Lemeshow P=0.45). Decision curve analysis determined clinical net benefit across thresholds of 15-40%. ML models, particularly random forest, accurately predicted DVT in patients with elderly hip fracture using routinely available admission data. Diabetes therefore emerged as a pivotal risk factor and its inclusion enhanced predictive accuracy. The present model holds promise for early risk stratification and individualized thromboprophylaxis, although rigorous external validation is warranted before any clinical application.

Correspondence to: Dr Qian Song, Department of Osteo-Internal Medicine, Tianjin University Tianjin Hospital, 406 Jiefang South Road, Hexi, Tianjin 300122, P.R. China
E-mail: songqian66882020@163.com

Dr Peng Tian, Department of Orthopedics, Tianjin University Tianjin Hospital, 406 Jiefang South Road, Hexi, Tianjin 300122, P.R. China
E-mail: tianpeng007@foxmail.com

*Contributed equally

Key words: deep vein thrombosis, hip fracture, machine learning, elderly, diabetes, prediction model, random forest

Introduction

Hip fractures represent a notable and growing global health burden among elderly populations, with annual incidence rates continuing to rise in tandem with aging demographics worldwide (1-3). Among the most common and devastating complications of hip fracture is lower extremity deep vein thrombosis (DVT), affecting 15-30% of patients in the absence of thromboprophylaxis (4). DVT not only increases morbidity and mortality but also marks an important juncture in functional independence, heralding a marked decline in mobility and quality of life (5). DVT can occur at any time after the fracture, often during the period of immobilization before surgery or in the postoperative phase and is associated with pulmonary embolism, post-thrombotic syndrome, prolonged hospitalization, increased mortality and notably higher healthcare expenditures (6).

The multifactorial pathogenesis of DVT involves complex interactions between patient vulnerability factors and precipitating insults, including fracture-associated immobility,

endothelial injury and hypercoagulability (7). This complexity presents notable challenges for accurate risk prediction using conventional statistical methods, such as multivariable logistic and Cox regression, which rely on manually specified interaction terms and assume linear predictor-outcome associations and therefore fail to automatically capture complex non-linear and interactive effects among multiple DVT risk factors (8).

Diabetes mellitus, a common comorbidity in elderly populations, creates a prothrombotic environment through multiple mechanisms: Endothelial dysfunction, platelet hyperreactivity, impaired fibrinolysis and elevated levels of coagulation factors (9,10). These metabolic alterations may synergistically interact with the hypercoagulable state induced by fracture and immobility to further elevate DVT risk, yet diabetes-specific quantitative and metabolic variables are often inadequately represented in traditional risk assessment tools such as the Caprini score. The Caprini score merely records a binary history of diabetes, excluding detailed metrics, including HbA1c, fasting glucose, diabetes duration, antidiabetic medication types and full lipid profiles that mediate thrombotic risk (11,12). In addition, lipid metabolism disorders frequently accompanying diabetes may contribute to thrombotic risk through effects on platelet function and coagulation pathways (13).

Traditional risk assessment tools, while clinically accessible, frequently demonstrate limited predictive performance in heterogeneous elderly populations due to their inability to adequately capture nonlinear associations and higher-order interactions among numerous risk factors (14,15). Furthermore, existing DVT prediction models often rely on postoperative or delayed clinical variables that may not be available at the time of admission, such as operation duration, intraoperative blood loss, postoperative drainage output and timing of first postoperative mobilization, limiting their utility for early risk stratification and prompt initiation of thromboprophylaxis (16,17).

The emerging application of machine learning (ML) in clinical prediction represents a paradigm shift in prognostic modeling, offering distinct advantages in handling complex, high-dimensional data and detecting subtle patterns that may elude conventional approaches (18,19). Recent advances have demonstrated the utility of ML across diverse biomedical domains, including molecular modeling for safer bioinformatics (20), fragment-based drug design with DNA binding evaluation (21), bioinformatics analysis of disease-associated biomarkers (22) and immunoproteomics for diagnostic and vaccine development (23). ML algorithms can automatically model intricate interactions among numerous clinical variables without requiring pre-specified hypotheses, potentially uncovering novel risk associations and improving predictive accuracy (18).

Current literature reveals a notable gap in comprehensive comparisons of numerous ML architectures specifically optimized for DVT prediction in patients with hip fracture, particularly using variables readily available at admission to facilitate early risk assessment (24). In addition, few studies have systematically examined the incremental value of diabetes-specific variables and lipid parameters or validated their importance through advanced interpretability techniques (25).

The present study aimed to address these research gaps by developing and validating numerous ML models for predicting DVT in elderly patients with hip fracture using clinical data available at the time of hospital admission. The present study specifically sought to compare the performance of diverse ML algorithms, including ensemble methods (random forest, XGBoost, LightGBM and CatBoost), support vector machine, logistic regression and decision tree, while simultaneously identifying and ranking key predictive features through advanced interpretability methods [including Shapley Additive exPlanations (SHAP) analysis], with a particular emphasis on diabetes and lipid parameters as key risk factors.

Materials and methods

Study design and population. A retrospective cohort study of patients aged ≥65 years who were admitted with hip fracture to a tertiary academic medical center (Tianjin Hospital, China) between January 2020 and December 2025 was conducted. All elderly patients with hip fracture during this period were consecutively screened for eligibility. Inclusion criteria: i) Age ≥65 years old, diagnosed with hip fracture through imaging (hip joint X-ray/CT), classified as femoral neck fracture or intertrochanteric fracture; ii) after admission, standardized and improved lower limb vein color Doppler ultrasound screening was conducted, and the diagnostic data for thrombosis was complete; iii) the clinical medical record data is complete, and all predictive variables such as demographics, underlying diseases, fracture injuries, test indicators, surgical anesthesia, perioperative medication and activity ability could be fully extracted; iv) hospitalization duration ≥5 days, no early automatic discharge or loss to follow-up, with a complete DVT screening imaging report; and v) accepted standard orthopedic surgical treatment: Proximal femoral nail anti-rotation (PFNA) internal fixation, artificial femoral head replacement and total hip replacement. Other exclusion criteria were as follows: i) A clear history of lower limb DVT and pulmonary embolism at the time of admission, and ultrasound findings of old venous thrombosis; ii) multiple severe traumas, pelvic fractures, spinal fractures, traumatic brain injury, hemorrhagic shock and continuous vascular active drug maintenance in ICU; iii) late-stage malignant tumors, maintenance hemodialysis, decompensated cirrhosis, severe coagulation dysfunction and autoimmune vasculitis; iv) long-term oral anticoagulant use (such as warfarin, rivaroxaban and apixaban) and long-term lower limb venous filter placement before admission; v) accepted standard orthopedic surgical treatment: PFNA internal fixation, artificial femoral head replacement or total hip replacement. Patients managed conservatively with long-term bed rest were excluded from the final analysis; vi) missing data for coagulation four items (prothrombin time, activated partial thromboplastin time, thrombin time and fibrinogen), D-dimer, fracture classification, surgical records and bilateral lower limb venous ultrasound reports; vii) long-term bed rest for >2 weeks before admission, paraplegia, lower limb disability and previous lower limb amputation surgery; viii) during hospitalization, due to death or automatic discharge, the standard lower limb venous ultrasound screening was not completed, and there were no reliable DVT judgment results; ix) non-traumatic hip fractures

(excluding pathological fractures caused by osteoporosis, and direct exclusion of bone marrow tumors and tumor metastasis fractures); and x) refusing thrombus screening and standardized anticoagulant prophylaxis after admission making it impossible to determine the true risk of DVT. After applying the exclusion criteria, a total of 400 patients (182 males and 218 females) were included in the final analysis. The present study was approved by the Institutional Review Board of Tianjin Hospital (Tianjin, China) Ethics Committee with a waiver for informed consent due to its retrospective nature.

Data collection and predictor variables. Data were systematically collected from electronic health records using a standardized protocol (26) by two independent reviewers, with disagreements resolved by consensus. The primary outcome was lower extremity DVT diagnosed by duplex ultrasound during hospitalization. DVT ascertainment followed a two-step process: Initial screening using International Classification of Diseases-9/10 codes for DVT [ICD-9 and ICD-10 diagnostic codes applied to ascertain DVT conform to standardized venous thromboembolism diagnostic coding criteria formulated by the World Health Organization (WHO) (27) and domestic hospital information coding specifications. ICD-9 coding inclusion criteria for DVT: Codes 451.11 and 451.19 were adopted, which correspond to phlebitis and thrombophlebitis of lower extremities, encompassing acute DVT of the femoral vein, popliteal vein, tibial vein and calf muscular veins; codes related to superficial phlebitis, chronic post-thrombotic syndrome and pulmonary embolism without concurrent lower-extremity DVT were excluded. ICD-10 coding inclusion criteria for DVT: Codes I82.4, I82.5, I82.6 and I82.8 were utilized, representing acute embolism and thrombosis of distal and proximal deep veins of the lower limb. Exclusion criteria for ICD-10 codes covered superficial venous thrombosis, chronic residual venous thrombotic lesions, thromboembolism of upper extremities, cerebral venous thrombosis and non-venous thrombotic disorders of visceral organs. All index admission diagnosis codes and secondary diagnosis codes meeting the above ICD-9/10 criteria were marked as DVT-positive flagged cases for subsequent standardized medical record abstraction], followed by structured chart review for all code-flagged patients and a 20% random sample of non-flagged patients. Both clinicians, blinded to predictor variables, independently reviewed radiology reports, progress notes and discharge summaries, with disagreements resolved by a third senior vascular specialist. Both screening and clinically indicated ultrasounds were included, with DVT defined as a new non-compressible venous segment or intraluminal filling defect.

A comprehensive set of variables were collected at the time of admission, encompassing demographic characteristics (including age, sex and BMI), comorbidities (hypertension, coronary artery disease, chronic kidney disease, diabetes mellitus with detailed specifications including type, duration and treatment modality, history of stroke or varicose veins), diabetes-specific measures [hemoglobin A1c (HbA1c), fasting glucose, insulin use and oral hypoglycemic agents], laboratory parameters [hemoglobin, platelet count, D-dimer, fibrinogen, albumin, creatinine, CRP and lipid profile including total cholesterol, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol and triglycerides], functional status (mobility status on admission categorized as independent ambulation, ambulatory with assistance or bedridden; Barthel Index assessed on admission), fracture characteristics (fracture type: Femoral neck vs. intertrochanteric) and a composite measure of comorbidity burden [Charlson Comorbidity Index (28)]. After data collection, preprocessing was performed following a standardized workflow (29) to ensure model development was leakage-free.

Data preprocessing and model development. A standardi zed ML development workflow was followed to ensure a rigorous and leakage-free evaluation. The dataset was first randomly split into a model development set (70%) and a hold-out internal validation set (30%) using stratified sampling based on the DVT outcome. All subsequent data-dependent steps, including handling of missing values, normalization and hyperparameter tuning, were conducted exclusively within the development set to prevent any information from the validation set influencing model development.

For variables with <20% missing values, multiple imputation by chained equations (MICE) with 10 iterations was applied to the training data. Convergence was assessed by visually inspecting trace plots of the mean and variance of imputed values across iterations; stable patterns without trends or strong autocorrelation were observed by iteration 5, with 10 iterations being conservatively used. The mean imputed values from the training set were then used to impute missing values in the test set. Variables with >20% missingness were excluded from analysis. Continuous variables were standardized using Z-score normalization, with the mean and SD calculated from the training set and applied to the test set. No feature selection was performed prior to modeling based on univariate statistical significance; all aforementioned variables listed were used as input features to allow the ML models to evaluate all potential associations, including complex interactions, thereby preventing information leakage from outcome-driven feature selection. Multicollinearity was assessed using variance inflation factors (VIF) for all predictor variables in the logistic regression model; all VIF values were <3, indicating no significant collinearity among the diabetes-associated variables or other predictors (Table SI). A total of seven ML algorithms were implemented: Logistic regression with L2 regularization as a baseline conventional model, decision tree for comparison as a single-tree architecture, random forest as an ensemble of decision trees with bagging, Extreme Gradient Boosting (XGBoost) as a gradient-boosted ensemble, LightGBM, CatBoost and support vector machine with a radial basis function kernel. For each algorithm, hyperparameter tuning was conducted strictly within the training set via grid search coupled with 10-fold cross-validation, optimizing for the area under the receiver operating characteristic (ROC) curve. The searched hyperparameter grids included: For random forest, number of trees (100, 300, 500), max depth (5, 10, 15) and min samples split (2, 5, 10); for XGBoost, learning rate (0.01, 0.1, 0.2), max depth (3, 6, 9) and n_estimators (100, 300); for LightGBM, learning rate (0.01, 0.05, 0.1), max depth (3, 5, 7), num_leaves (15, 31, 63) and min_child_samples (10, 20, 30); for CatBoost, iterations (100, 300, 500), learning rate (0.01, 0.05, 0.1), depth (3, 5, 7) and l2_leaf_reg (1, 3, 5); for support

vector machine, C (0.1, 1, 10) and γ ('scale', 'auto', 0.01, 0.1); for logistic regression, C (0.01, 0.1, 1, 10); and for decision tree, max depth (3, 5, 10, 15) and min samples split (2, 5, 10). The model configuration achieving the highest mean cross-validated area under the curve (AUC) was selected for each algorithm. The final model for each algorithm was produced by retraining the optimal hyperparameter configuration on the entire preprocessed training set and then applied only once to the preprocessed hold-out internal validation set for final, unbiased evaluation. The search ranges for hyperparameters were determined based on preliminary experiments [including i) trial runs testing different iteration numbers (5/10/15) for MICE multiple imputation to identify sufficient convergence iterations; ii) testing of varying Z-score normalization strategies to confirm leakage-free scaling logic; iii) pre-screening of hyperparameter value ranges for all seven machine learning algorithms via small subset training samples, eliminating obviously underfitting or overfitting parameter combinations; iv) VIF multicollinearity pre-tests on raw feature sets to exclude highly collinear variable combinations beforehand; and v) pilot training with partial cohort data to compare rough AUC performance boundaries of each algorithm, laying the range basis for formal grid search], common configurations reported (30,31) in the literature and the need to cover a spectrum from simple to complex models for each algorithm. After developing and tuning all seven models, their performance was evaluated on the hold-out validation set using a comprehensive set of metrics.

Model evaluation and validation. Model performance was assessed on the internal validation set using the area under the ROC curve with 95% CIs calculated by DeLong's method, along with accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F1-score. DeLong's tests were also performed to compare AUCs between models, particularly random forest vs. the others. Calibration was evaluated using calibration plots, calibration slope and intercept, the Hosmer-Lemeshow goodness-of-fit test and the Brier score. Clinical utility was assessed using decision curve analysis across probability thresholds from 0-0.5 and a clinical impact curve to illustrate the number of patients classified as high-risk and the number of true positives at various risk thresholds.

To enhance interpretability, feature importance was calculated for tree-based models as mean decrease in impurity. In addition, SHAP was applied to the best-performing model to generate summary plots of global feature importance and direction of effect, dependence plots for key variables (including diabetes-associated and lipid parameters) to visualize nonlinear associations and interactions and to quantify the contribution of these variables to overall predictions.

The performance of the random forest model was evaluated in prespecified subgroups, including diabetic vs. non-diabetic patients, age <80 vs. ≥80 years, men vs. women, BMI <25 vs. ≥25 kg/m^2 and mobility status on admission (independent vs. assisted or bedridden). AUC values with 95% CIs were calculated for each subgroup.

To assess robustness, a number of sensitivity analyses were conducted: Complete-case analysis excluding patients with any missing data, repeating model development with five different random seeds for train-test split, using mean imputation instead of MICE, and comparing the primary random forest model to one trained only on variables with P<0.05 from univariate analysis.

Reporting guidelines. Analysis was conducted and the present study results were reported in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-Artificial Intelligence statement (32).

Statistical analysis. Data distribution was tested using the Shapiro-Wilk test. Patient characteristics are described using mean ± SD or median with interquartile range for continuous variables and frequency with percentages for categorical variables. Independent t-tests or Mann-Whitney U tests were applied for continuous variables with normal or non-normal distribution, respectively. Categorical variables were analyzed using the Pearson χ^2 test or Fisher's exact test as appropriate.

Univariate logistic regression analysis was performed to assess the unadjusted associations between individual variables measured at admission and DVT, providing a conventional clinical summary of potential risk factors. All variables, including demographic characteristics, comorbidities, diabetes-specific laboratory markers, routine laboratory parameters, functional status indicators, fracture classification and Charlson Comorbidity Index score (detailed list: age, sex, BMI, hypertension, coronary artery disease, chronic kidney disease, diabetes mellitus, stroke history, varicose veins, HbA1c, fasting glucose, insulin use, oral hypoglycemic agents, hemoglobin, platelet count, D-dimer, fibrinogen, albumin, creatinine, CRP, lipid profile, mobility status, Barthel Index, fracture type, Charlson Comorbidity Index) were first analyzed via univariate logistic regression. All of these variables, regardless of univariate statistical significance, were adopted as input features for subsequent machine learning modeling.

All statistical analyses were conducted using SPSS software (version 26.0; IBM Corp) and Python (version 3.9.16; Python Software Foundation) with libraries including scikit-learn 1.2.2, XGBoost 1.7.5, LightGBM 3.3.5, CatBoost 1.2, SHAP 0.41.0, matplotlib 3.7.1 and pandas 1.5.3. MICE was implemented using the IterativeImputer class from sklearn.impute with max_iter=10, tol=1e-3 and convergence monitored via the change in parameter estimates between successive iterations. VIF calculations were performed using the 'variance_inflation_factor' function from 'statsmodels.stats.outliers_influence' in Python. Two-tailed P<0.05 was considered to indicate a statistically significant difference for baseline comparisons and univariate analysis.

Results

Baseline characteristics and DVT incidence. A total of 400 consecutive elderly patients admitted with hip fracture who met the inclusion criteria were included the final analytical cohort. The overall incidence of lower extremity DVT during hospitalization was 16.5% (66/400 patients). As detailed in Table I, baseline characteristics differed significantly between patients who developed DVT and those who did not. Patients in the DVT group were significantly older (83.1±7.3 vs. 78.4±8.6 years; P<0.001) and had a lower BMI (22.8±3.3 vs. 24.1±3.7 kg/m^2; P=0.012). The prevalence of diabetes

Table I. Baseline characteristics of study population stratified by DVT status.

Characteristic	Total (n=400)	DVT group (n=66)	Non-DVT group (n=334)	P-value
Demographics				
Mean age ± SD, years	79.20±8.50	83.10±7.30	78.40±8.60	<0.001
Men, n (%)	176 (44.0)	26 (39.4)	150 (44.9)	0.421
Mean BMI ± SD, kg/m^2	23.90±3.60	22.80±3.30	24.10±3.70	0.012
Comorbidities, n (%)				
Hypertension	276 (69.0)	50 (75.8)	226 (67.7)	0.198
Coronary artery disease	148 (37.0)	32 (48.5)	116 (34.7)	0.032
Chronic kidney disease	64 (16.0)	18 (27.3)	46 (13.8)	0.006
Diabetes mellitus	130 (32.5)	34 (51.5)	96 (28.7)	<0.001
Type 1 diabetes	10 (2.5)	3 (4.5)	7 (2.1)	0.215
Type 2 diabetes	120 (30.0)	31 (47.0)	89 (26.6)	<0.001
Mean diabetes duration ± SD, years	8.2±5.4	11.8±6.3	7.4±4.9	<0.001
Insulin use	44 (11.0)	15 (22.7)	29 (8.7)	<0.001
Oral hypoglycemic agents	86 (21.5)	19 (28.8)	67 (20.1)	0.124
History of stroke	58 (14.5)	16 (24.2)	42 (12.6)	0.012
Varicose veins	40 (10.0)	12 (18.2)	28 (8.4)	0.012
Heart failure	54 (13.5)	15 (22.7)	39 (11.7)	0.014
Atrial fibrillation	48 (12.0)	14 (21.2)	34 (10.2)	0.010
Charlson Comorbidity Index, mean ± SD	4.40±2.20	5.40±2.40	4.20±2.10	<0.001
Laboratory parameters, mean ± SD				
HbA1c, %	6.70±1.30	7.60±1.50	6.50±1.20	<0.001
Fasting glucose, mmol/l	6.75±1.92	7.68±2.18	6.58±1.82	<0.001
Hemoglobin, g/l	123.80±18.20	113.60±17.50	125.80±18.00	<0.001
Platelet count, x10^9/l	196.50±54.80	208.20±59.40	194.30±53.60	0.058
D-dimer, ng/ml	6120±4580	8920±5380	5560±4280	<0.001
Fibrinogen, g/l	3.92±1.18	4.58±1.32	3.79±1.12	<0.001
Albumin, g/l	36.50±5.00	31.80±4.80	37.40±4.90	<0.001
Hypoalbuminemia <35 g/l, n (%)	128 (32.0)	38 (57.6)	90 (26.9)	<0.001
Creatinine, μmol/l	79.80±26.50	89.20±29.60	77.90±25.40	0.002
CRP, mg/l	26.80±19.50	36.50±22.80	24.90±18.40	<0.001
Total cholesterol, mmol/l	4.78±1.15	5.22±1.22	4.69±1.12	<0.001
LDL cholesterol, mmol/l	2.92±0.88	3.38±0.95	2.83±0.84	<0.001
HDL cholesterol, mmol/l	1.30±0.34	1.14±0.31	1.33±0.35	<0.001
Triglycerides, mmol/l	1.68±0.75	2.02±0.86	1.61±0.71	<0.001
Functional status				
Mobility on admission, n (%)				<0.001
Independent ambulation	158 (39.5)	14 (21.2)	144 (43.1)	
Ambulatory with assistance	155 (38.8)	28 (42.4)	127 (38.0)	
Bedridden	87 (21.8)	24 (36.4)	63 (18.9)	
Barthel Index on admission, mean ± SD	61.20±17.20	46.50±15.80	64.10±16.80	<0.001
Fracture characteristics				
Fracture type, n (%)				0.142
Femoral neck	212 (53.0)	31 (47.0)	181 (54.2)	
Intertrochanteric	188 (47.0)	35 (53.0)	153 (45.8)	
Mean time from injury to admission ± SD, days	1.80±1.50	2.40±1.80	1.70±1.40	0.002
Mean time from admission to surgery ± SD, days	3.50±2.20	4.60±2.60	3.30±2.10	<0.001

Data are presented as mean ± SD or n (%). P-values were calculated using independent t-tests for continuous variables and χ^2 tests for categorical variables. DVT, deep vein thrombosis; LDL, low-density lipoprotein; HDL, high-density lipoprotein; HbA1c, hemoglobin A1c.

mellitus was significantly higher in the DVT group (51.5 vs. 28.7%; P<0.001), with longer diabetes duration (11.8±6.3 vs. 7.4±4.9 years; P<0.001) and higher rates of insulin use (22.7 vs. 8.7%; P<0.001). Other comorbidities significantly associated with DVT included coronary artery disease (48.5 vs. 34.7%; P=0.032), chronic kidney disease (27.3 vs. 13.8%; P=0.006), history of stroke (24.2 vs. 12.6%; P=0.012), varicose veins (18.2 vs. 8.4%; P=0.012), heart failure (22.7 vs. 11.7%; P=0.014) and atrial fibrillation (21.2 vs. 10.2%; P=0.010). The Charlson Comorbidity Index was also significantly higher in the DVT group (5.4±2.4 vs. 4.2±2.1; P<0.001).

Laboratory parameters revealed a distinct prothrombotic and metabolic profile among patients who developed DVT. These patients exhibited significantly higher HbA1c levels (7.6±1.5 vs. 6.5±1.2%; P<0.001) and fasting glucose (7.68±2.18 vs. 6.58±1.82 mmol/l; P<0.001). Markers of thrombosis and inflammation were also significantly higher, including D-dimer (8920±5380 vs. 5560±4280 ng/ml; P<0.001) and fibrinogen (4.58±1.32 vs. 3.79±1.12 g/l; P<0.001). In addition, CRP was significantly higher in the DVT group (36.5±22.8 vs. 24.9±18.4 mg/l; P<0.001). Nutritional status was significantly poorer in the DVT group, as evidenced by lower albumin levels (31.8±4.8 vs. 37.4±4.9 g/l; P<0.001) and a higher prevalence of hypoalbuminemia (57.6 vs. 26.9%; P<0.001). Hemoglobin was also significantly lower (113.6±17.5 vs. 125.8±18.0 g/l; P<0.001). Regarding lipid profiles, patients with DVT exhibited significantly higher total cholesterol (5.22±1.22 vs. 4.69±1.12 mmol/l; P<0.001), LDL cholesterol (3.38±0.95 vs. 2.83±0.84 mmol/l; P<0.001) and triglycerides (2.02±0.86 vs. 1.61±0.71 mmol/l; P<0.001), while HDL cholesterol was significantly lower (1.14±0.31 vs. 1.33±0.35 mmol/l; P<0.001).

Functional status on admission was significantly impaired in the DVT group compared with that in the non-DVT group, with a lower Barthel Index (46.5±15.8 vs. 64.1±16.8; P<0.001) and a higher proportion of patients requiring assistance or being bedridden (78.8 vs. 56.9%; P<0.001). Process-associated factors also differed significantly: Patients who developed DVT exhibited a longer time from injury to admission (2.4±1.8 vs. 1.7±1.4 days; P=0.002) and longer time from admission to surgery (4.6±2.6 vs. 3.3±2.1 days; P<0.001). Fracture type was similarly distributed between groups (P=0.142). To further quantify the strength of association for each risk factor, univariate logistic regression analysis was performed.

Univariate analysis of risk factors for DVT. Univariate logistic regression analysis revealed a wide spectrum of significant predictors spanning demographic, comorbidity, laboratory, functional and process-associated domains, as detailed in Table II. Among the strongest identified risk factors were diabetes mellitus [odds ratio (OR)=2.64; 95% CI: 1.56-4.48; P<0.001], with each additional year of diabetes duration conferring a 10% increase in DVT risk (OR=1.10; 95% CI: 1.05-1.15; P<0.001). Insulin use was associated with a >2-fold elevated risk (OR=2.38; 95% CI: 1.38-4.12; P=0.002). Hypoalbuminemia emerged as one of the most potent predictors, with a nearly 4-fold increased odds of DVT (OR=3.58; 95% CI: 2.08-6.16; P<0.001), reinforced by the continuous protective effect of albumin (OR=0.84 per g/l; 95% CI: 0.79-0.89; P<0.001).

Markers of thrombosis and inflammation demonstrated strong associations: D-dimer (OR=1.28 per 1,000 ng/ml; 95% CI: 1.18-1.39; P<0.001), fibrinogen (OR=1.62 per g/l; 95% CI: 1.32-1.99; P<0.001) and CRP (OR=1.22 per 10 mg/l; 95% CI: 1.08-1.38; P<0.001) were all significantly associated with increased DVT risk. Lipid parameters also contributed notably: LDL cholesterol (OR=1.58 per mmol/l; 95% CI: 1.26-1.98; P<0.001), triglycerides (OR=1.58 per mmol/l; 95% CI: 1.24-2.01; P<0.001) and total cholesterol (OR=1.38 per mmol/l; 95% CI: 1.12-1.70; P<0.001) were significant risk factors, while HDL cholesterol demonstrated a protective effect (OR=0.58 per mmol/l; 95% CI: 0.42-0.80; P<0.001).

Functional status was highly predictive: Compared with independently mobile patients, those requiring assistance exhibited a >2-fold increased risk (OR=2.28; 95% CI: 1.22-4.26; P=0.010), and bedridden patients showed a ~4-fold increased risk (OR=3.92; 95% CI: 1.98-7.76; P<0.001). Each 10-point decrement in Barthel Index was associated with a 38% increase in DVT risk (OR=0.62; 95% CI: 0.51-0.75; P<0.001). Process-associated factors also served key roles: Each additional day from injury to admission increased risk by 25% (OR=1.25; 95% CI: 1.08-1.45; P=0.002) and each additional day from admission to surgery increased risk by 18% (OR=1.18; 95% CI: 1.08-1.29; P<0.001). While univariate analysis identifies individual risk factors, ML models can capture complex interactions among multiple predictors. Therefore seven different algorithms were evaluated for DVT prediction.

Comparative performance of ML algorithms. Predictive performances of seven distinct ML classifiers for DVT were evaluated on the internal validation set, with results systematically summarized in Table III and graphically presented in Fig. 1. Fig. 1A-G display the individual ROC curves for each model with their corresponding area AUC values. Fig. 1H overlays the ROC curves of all seven models, allowing direct visual comparison. The diagonal dashed line in each panel represents the reference line of no discrimination.

Among all evaluated algorithms, the ensemble-based random forest model demonstrated improved performance, achieving an AUC of 0.92 (95% CI: 0.87-0.96). This high discriminative ability was complemented by balanced and robust performance across all other key metrics, including an accuracy of 0.87, sensitivity of 0.85, specificity of 0.88, positive predictive value of 0.80, negative predictive value of 0.91 and an F1-score of 0.84.

Other gradient boosting methods also exhibited strong predictive capability: LightGBM achieved an AUC of 0.91 (95% CI: 0.86-0.95), while XGBoost and CatBoost both attained AUCs of 0.90 (95% CI: 0.85-0.94). The support vector machine classifier achieved moderate performance with an AUC of 0.87 (95% CI: 0.82-0.92). By contrast, both the conventional logistic regression and the simpler Decision Tree classifiers showed comparatively limited performance, with AUC values of 0.84 (95% CI: 0.78-0.89) and 0.81 (95% CI: 0.75-0.87), respectively.

DeLong's test comparisons revealed that the random forest AUC was significantly higher compared with logistic regression (P=0.006) and Decision Tree (P<0.001), but the differences with XGBoost (P=0.220), LightGBM (P=0.280)

Table II. Univariate logistic regression analysis of risk factors for deep vein thrombosis.

Variable	OR	95% CI	P-value
Demographics			
Age, per year	1.07	1.04-1.10	<0.001
Men (vs. women)	0.82	0.49-1.37	0.421
BMI, per kg/m^2	0.95	0.91-0.99	0.012
Comorbidities			
Hypertension	1.48	0.82-2.68	0.198
Coronary artery disease	1.78	1.06-3.01	0.032
Chronic kidney disease	2.35	1.28-4.32	0.006
Diabetes mellitus	2.64	1.56-4.48	<0.001
Diabetes duration, per year	1.10	1.05-1.15	<0.001
Insulin use	2.38	1.38-4.12	0.002
Oral hypoglycemic agents	1.52	0.89-2.58	0.124
History of stroke	2.15	1.18-3.92	0.012
Varicose veins	2.38	1.22-4.65	0.012
Heart failure	2.18	1.16-4.09	0.014
Atrial fibrillation	2.38	1.22-4.65	0.010
Charlson Comorbidity Index, per point	1.24	1.11-1.39	<0.001
Laboratory parameters			
HbA1c, per 1%	1.42	1.22-1.65	<0.001
Fasting glucose, per mmol/l	1.22	1.10-1.35	<0.001
Hemoglobin, per 10 g/l	0.82	0.74-0.91	<0.001
Platelet count, per 50x10^9/l	1.12	0.96-1.31	0.058
D-dimer, per 1,000 ng/ml	1.28	1.18-1.39	<0.001
Fibrinogen, per g/l	1.62	1.32-1.99	<0.001
Albumin, per g/l	0.84	0.79-0.89	<0.001
Hypoalbuminemia, <35 g/l	3.58	2.08-6.16	<0.001
Creatinine, per 10 μmol/l	1.11	1.03-1.20	0.002
CRP, per 10 mg/l	1.22	1.08-1.38	<0.001
Total cholesterol, per mmol/l	1.38	1.12-1.70	<0.001
LDL cholesterol, per mmol/l	1.58	1.26-1.98	<0.001
HDL cholesterol, per mmol/l	0.58	0.42-0.80	<0.001
Triglycerides, per mmol/l	1.58	1.24-2.01	<0.001
Functional status			
Mobility on admission	1.35	1.02-1.68	<0.001
Independent ambulation	1.56	1.21-1.99	<0.001
Ambulatory with assistance	2.28	1.22-4.26	0.010
Bedridden	3.92	1.98-7.76	<0.001
Barthel Index, per 10 points	0.62	0.51-0.75	<0.001
Fracture or process factors			
Intertrochanteric, vs. femoral neck	1.32	0.79-2.22	0.142
Time from injury to admission, per day	1.25	1.08-1.45	0.002
Time from admission to surgery, per day	1.18	1.08-1.29	<0.001

OR, odds ratio; LDL, low-density lipoprotein; HDL, high-density lipoprotein; HbA1c, hemoglobin A1c.

and CatBoost (P=0.240) were not statistically significant. The random forest model also demonstrated marked calibration on the internal validation set (Fig. 2), with a calibration slope of 1.02 (95% CI: 0.95-1.09), an intercept of -0.02, a Brier score of 0.09 and a non-significant Hosmer-Lemeshow test (P=0.450), indicating that its predicted risks were reliable across the probability spectrum. Having established the improved performance of the random forest model, the present study further examined which features drove its predictions using SHAP analysis.

Table III. Performance comparison of machine learning models on the internal validation set for deep vein thrombosis prediction.

Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1-score	Brier score
Random forest	0.92 (0.87-0.96)	0.87	0.85	0.88	0.80	0.91	0.84	0.09
XGBoost	0.90 (0.85-0.94)	0.86	0.84	0.87	0.78	0.90	0.83	0.10
LightGBM	0.91 (0.86-0.95)	0.87	0.85	0.88	0.79	0.91	0.84	0.09
CatBoost	0.90 (0.85-0.94)	0.86	0.84	0.87	0.78	0.90	0.83	0.10
Support vector machine	0.87 (0.82-0.92)	0.83	0.81	0.84	0.74	0.88	0.79	0.12
Logistic regression	0.84 (0.78-0.89)	0.81	0.79	0.82	0.71	0.87	0.76	0.13
Decision tree	0.81 (0.75-0.87)	0.79	0.77	0.80	0.68	0.85	0.73	0.15

DeLong's test comparisons: Random forest vs. logistic regression (P=0.006); random forest vs. decision tree (P<0.001); random forest vs. XGBoost (P=0.22); random forest vs. LightGBM (P=0.28); random forest vs. CatBoost (P=0.24); and random forest vs. support vector machine (P=0.03). AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; XGBoost, Extreme Gradient Boosting.

Variable importance and SHAP analysis. Comprehensive feature importance analysis across the seven ML approaches revealed notable consistency in identifying core predictive factors, as quantitatively detailed in Table IV. Random forest feature importance based on mean decrease in impurity identified HbA1c as the most important predictor (normalized score: 100), followed by D-dimer (normalized score: 96), age (normalized score: 90), diabetes duration (normalized score: 88), albumin (normalized score: 86), time from admission to surgery (normalized score: 82), LDL cholesterol (normalized score: 78), immobility on admission (normalized score: 76) and triglycerides (normalized score: 72). SHAP analysis further determined these rankings and provided additional insights into the direction and nature of effects (Fig. 3).

HbA1c showed a strong positive association with DVT risk, with SHAP values increasing sharply at levels above 7.0%, suggesting a potential threshold effect for glycemic control. Diabetes duration exhibited a near-linear positive association with DVT risk. Albumin demonstrated a negative association, with risk increasing notably when albumin fell below 35 g/l. Age exhibited a gradual positive association, with accelerating risk after age 80 years (Fig. 3A).

Among thrombotic markers, D-dimer displayed a nonlinear effect, with risk increasing steeply up to ~8,000 ng/ml before plateauing at high levels. Fibrinogen exhibited a more linear positive association. Lipid parameters revealed important patterns: LDL cholesterol and triglycerides exhibited positive associations with DVT risk, while HDL cholesterol demonstrated a consistent protective effect.

SHAP dependence plots revealed notable interactions among key predictors (Fig. 3B and C). The effect of HbA1c was modified by age, such that the HbA1c-associated risk increase was more notable in patients aged ≥80 years. Diabetes duration interacted with D-dimer, with patients exhibiting a long diabetes duration and elevated D-dimer experiencing a disproportionately high risk. The effect of LDL cholesterol was potentiated in patients with elevated triglycerides. Notably, diabetes-associated variables (HbA1c, diabetes duration, fasting glucose and insulin use) collectively accounted for ~30% of total feature importance in the random forest model, while lipid parameters collectively contributed an additional 18%, underscoring the pivotal role of metabolic factors in DVT pathogenesis.

Subgroup analysis. Random forest models maintained robust and consistent discriminative ability across all prespecified subgroups, as detailed in Table V and visualized in Fig. 4. In diabetic patients, the model achieved an AUC of 0.94 (95% CI: 0.89-0.98) compared with 0.90 (95% CI: 0.85-0.94) in non-diabetic patients. For patients aged ≥80 years, the AUC was 0.92 (95% CI: 0.87-0.96) vs. 0.91 (95% CI: 0.86-0.95) for those <80 years. Performance was similarly robust across sexes, with AUCs of 0.91 for men and 0.92 for women. Of note, the AUC of 0.95 in the HbA1c ≥7.0% subgroup, while promising, should be interpreted with caution given the moderate sample size (n=108; events=28). These findings require demonstration in larger prospective cohorts.

Patients with lower BMI (<25 kg/m^2) exhibited a slightly higher model performance (AUC of 0.93) compared with those with a higher BMI (AUC of 0.90). The model

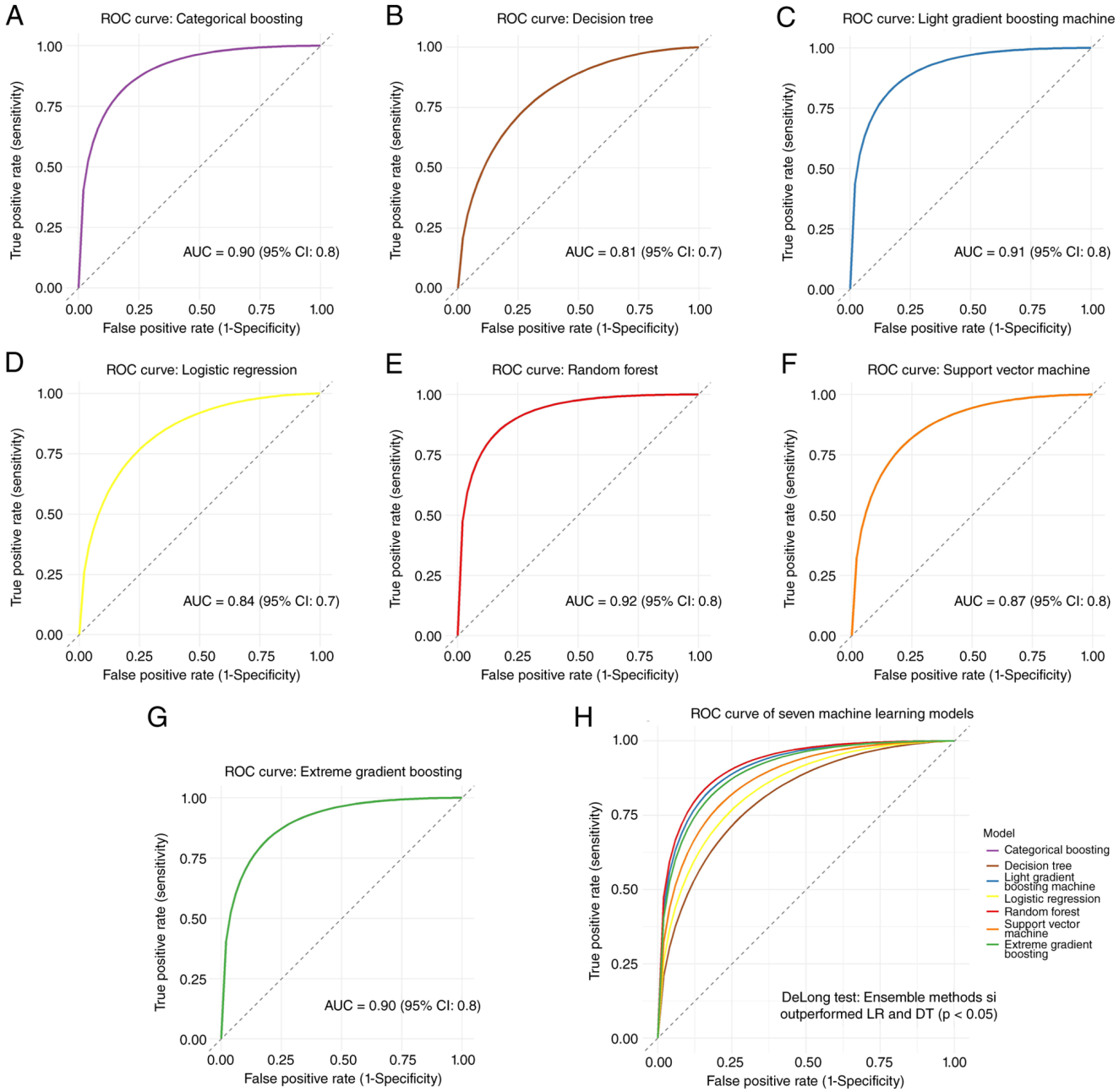

Figure 1. ROC curves of seven ML models. ROC curves for (A) CatBoost, (B) decision tree, (C) LightGBM, (D) logistic regression, (E) random forest, (F) SVM and (G) XGBoost, with AUC values as indicated. (H) Overlay of the ROC curves of all seven models, with corresponding AUC values listed: CatBoost, 0.90; decision tree, 0.81; LightGBM, 0.91; logistic regression, 0.84; random forest, 0.92; SVM, 0.87; XGBoost, 0.90. The diagonal dashed line represents the reference line of no discrimination. ROC, receiver operating characteristic; ML, machine learning; AUC, area under curve; SVM, support vector machine; XGBoost, Extreme Gradient Boosting.

demonstrated good discrimination in patients with impaired mobility on admission (AUC of 0.93) compared with independently mobile patients (AUC of 0.88). Notably, among patients with HbA1c ≥7.0%, the model achieved an AUC of 0.95, markedly higher compared with those with improved glycemic control (AUC of 0.89). Similarly, patients with elevated LDL cholesterol (≥3.0 mmol/l) exhibited a higher model performance (AUC of 0.93) compared with those with lower LDL (AUC of 0.90). The model also performed well in patients with prolonged preoperative waiting time (>3 days), achieving an AUC of 0.93 vs. 0.89 in those with shorter waiting times.

Sensitivity analysis. Multiple sensitivity analyses demonstrated the robustness of the primary findings, as detailed in Table VI. Complete-case analysis, excluding patients with any missing data (n=352), yielded a random forest AUC of 0.91 (95% CI: 0.86-0.95), nearly identical to the primary analysis using imputed data. A total of five different random seeds for train-test split produced random forest AUCs ranging from 0.90-0.93, with a mean of 0.92 and SD of 0.01, indicating good stability. Using mean imputation instead of MICE resulted in an AUC of 0.91 (95% CI: 0.86-0.95), comparable with the primary analysis. The model trained only on variables with P<0.05 from univariate analysis achieved an AUC of 0.91,

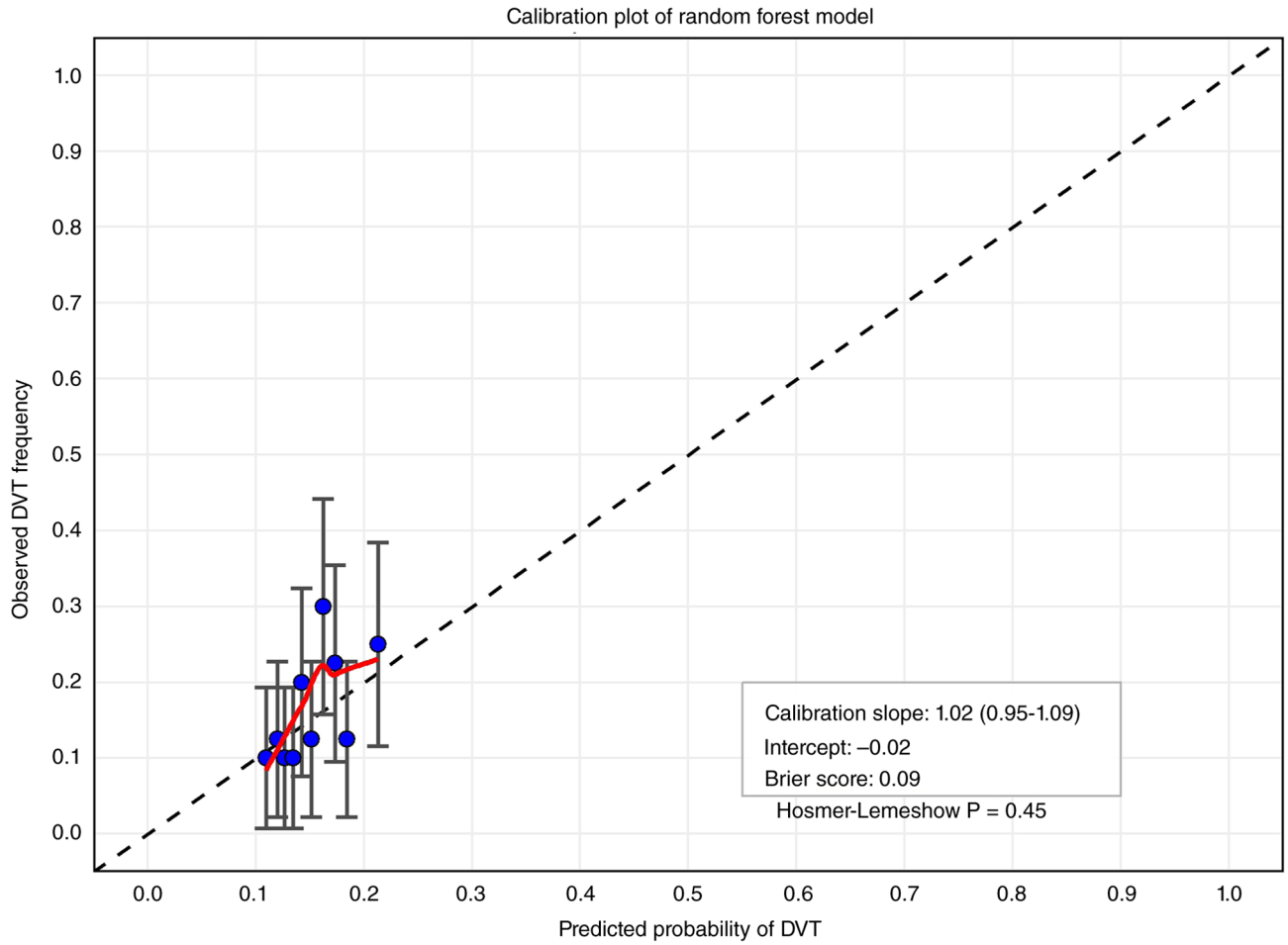

Figure 2. Calibration plot of the random forest model. Calibration metrics for the random forest model: A calibration slope of 1.02 (95% CI: 0.95-1.09), intercept of -0.02, Brier score of 0.09 and a Hosmer-Lemeshow P=0.45, indicating good calibration. The x-axis represents the predicted probability of DVT and the y-axis the observed DVT frequency. DVT, deep vein thrombosis.

confirming no performance degradation caused by retaining all potential predictors in the full model.

To assess the incremental value of specific variable categories, additional sensitivity analyses were performed. Excluding lipid parameters from the model reduced the AUC to 0.88 (95% CI: 0.83-0.92), demonstrating the marked contribution of lipid profile to prediction accuracy. Excluding diabetes-specific variables (HbA1c, diabetes duration, fasting glucose and insulin use) resulted in a more notable reduction to 0.85 (95% CI: 0.80-0.90), underscoring the pivotal role of diabetes-associated factors in DVT prediction. A 10-fold cross-validation on the training set yielded a mean AUC of 0.91 (95% CI: 0.87-0.95), consistent with the hold-out validation results.

Clinical utility analysis. Decision curve analysis (Fig. 5) demonstrated that the random forest model provided positive net benefit across a wide range of clinically relevant threshold probabilities from 10-45%, outperforming both the treat-all and treat-none strategies as well as the other ML models. At a risk threshold of 20%, the net benefit was 0.13, equivalent to correctly identifying 13 additional true DVT cases per 100 patients without increasing false positives.

The clinical impact curve (Fig. 6) illustrated the practical implications of model implementation: At a 20% risk threshold, ~28% of patients would be classified as high-risk, and these high-risk patients would account for 86% of all DVT cases (sensitivity). The positive predictive value at this threshold was 50.7%, indicating that among patients identified as high-risk, ~50% would actually develop DVT. This translates to a number needed to treat (NNT) of ~2.0, meaning that for every 2 patients classified as high-risk and receiving targeted prophylaxis, one DVT event could potentially be prevented. This suggested favorable cost-effectiveness for targeted thromboprophylaxis programs, while avoiding unnecessary intervention in 72% of low-risk patients.

Proposed clinical decision pathway. Based on the present exploratory findings, a hypothesis-generating clinical decision pathway may be proposed for risk stratification and prevention of DVT in elderly patients with hip fracture, as illustrated in Fig. 7. This conceptual framework requires prospective validation before clinical implementation. Upon admission, key clinical data including diabetes-specific indicators (HbA1c, diabetes duration and insulin use), age, lipid profile, D-dimer, albumin, mobility status and time from injury/admission are collected. These variables are input into the validated random

Table IV. Feature importance comparison across machine learning algorithms.

Variable	Random forest	XGBoost	LightGBM	CatBoost	SVM	Logistic regression	Decision tree	Mean ± SD
HbA1c	100	98	99	97	92	85	89	94.3±5.6
D-dimer	96	95	94	93	89	83	87	91.0±4.8
Age	90	88	89	87	85	81	83	86.1±3.3
Diabetes duration	88	89	88	86	79	75	77	83.1±5.8
Albumin	86	84	85	83	80	78	81	82.4±2.9
Time, admission to surgery	82	80	81	79	74	72	76	77.7±3.9
LDL cholesterol	78	76	79	77	70	67	66	73.3±5.4
Immobility on admission	76	78	77	75	71	69	73	74.1±3.4
Triglycerides	72	70	74	71	66	64	62	68.4±4.6
Fibrinogen	68	66	67	65	62	60	61	64.1±3.1
Fasting glucose	66	68	65	64	60	58	59	62.9±3.8
BMI	62	60	61	59	56	54	57	58.4±2.9
Time, injury to admission	60	58	59	57	54	52	55	56.4±2.9
HDL cholesterol	58	56	57	55	52	50	51	54.1±3.1
CRP	54	52	53	51	48	46	47	50.1±3.1
Chronic kidney disease	50	48	49	47	44	45	42	46.4±2.9
Atrial fibrillation	46	44	45	43	40	41	39	42.6±2.6
Varicose veins	42	40	41	39	36	37	35	38.6±2.6
Heart failure	38	36	37	35	33	34	32	35.0±2.2
Barthel Index	34	32	33	31	30	29	31	31.4±1.7
Creatinine	30	29	28	27	26	25	27	27.4±1.8
Hemoglobin	26	25	24	23	22	23	22	23.6±1.5
History of stroke	22	21	20	19	18	19	20	19.9±1.3
Total cholesterol	18	17	16	15	14	16	15	15.9±1.3
Coronary artery disease	14	13	12	11	12	13	12	12.4±1.0
Hypertension	10	9	8	7	10	11	9	9.1±1.3
Fracture type	8	7	6	5	8	9	7	7.1±1.3
Sex	6	5	4	3	6	7	5	5.1±1.3

Importance scores are normalized to a 0-100 scale for comparability across algorithms. For tree-based models (random forest, XGBoost, LightGBM, CatBoost and decision tree), importance was based on mean decrease in impurity. For logistic regression, importance was based on absolute coefficient values. For support vector machine, importance was based on permutation importance. LDL, low-density lipoprotein; HDL, high-density lipoprotein; XGBoost, Extreme Gradient Boosting; HbA1c, hemoglobin A1c.

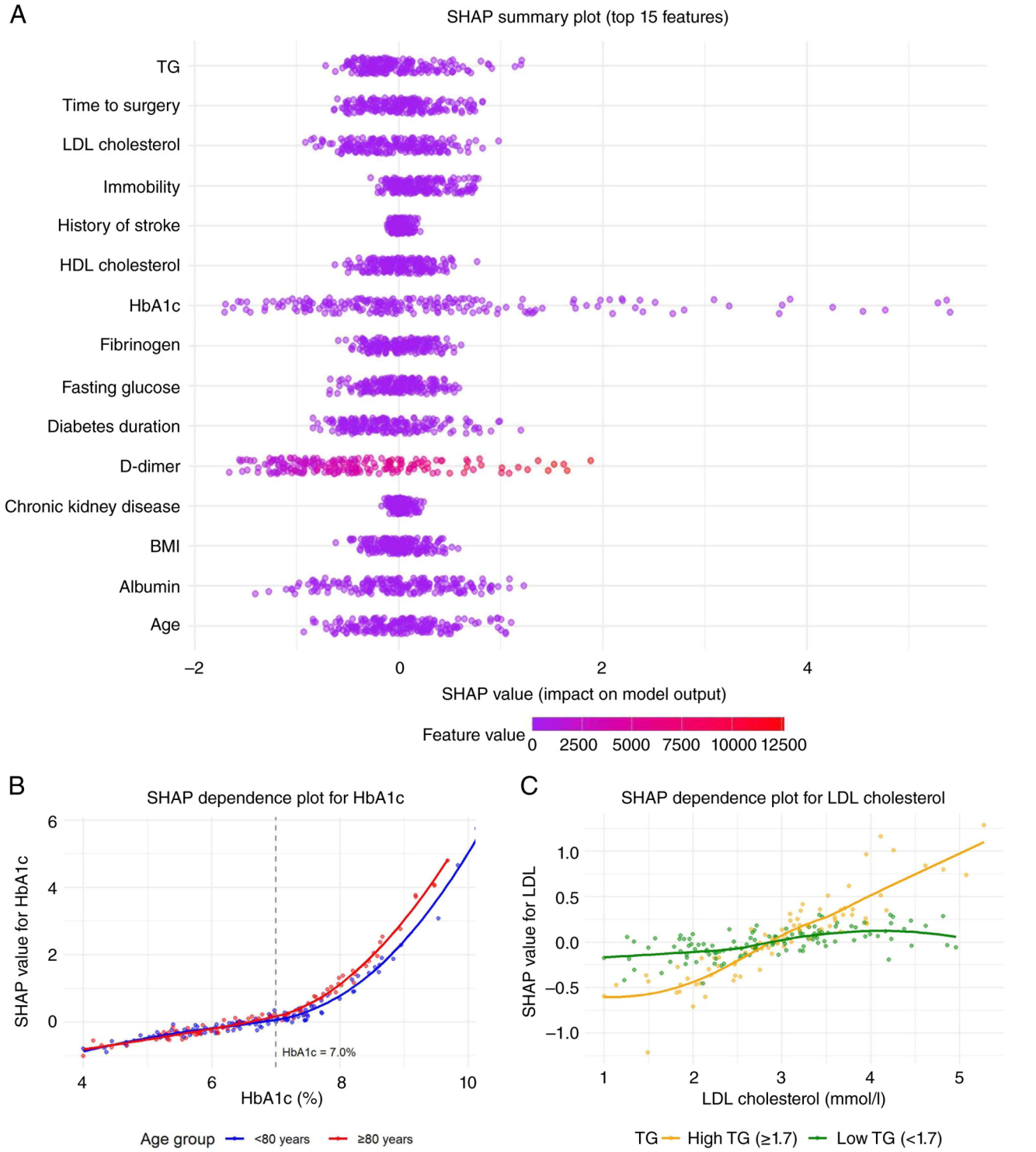

Figure 3. SHAP analysis of the random forest model. (A) SHAP summary plot listing the top 15 features ranked by mean absolute SHAP value. The color bar indicates feature value (red indicates high and blue indicates low). Positive SHAP values correspond to increased DVT risk. The features shown (from top to bottom) are: TG, time to surgery, LDL cholestrol, immobility, history of stroke, HDL cholestrol, HbA1c, Fibrinogen, fasting glucose, diabetes duration, D-dimer, chronic kidney disease, BMI, albumin and age. (B) SHAP dependence plot for HbA1c. Points are colored by age group (blue indicates <80 years and red indicates ≥80 years). The vertical dashed line marks the threshold HbA1c=7.0%. (C) SHAP dependence plot for LDL cholesterol. Points are colored by TG level (red indicates low TG <1.7 mmol/l and green indicates high TG ≥1.7 mmol/l). SHAP, Shapley Additive exPlanations; DVT, deep vein thrombosis; LDL, low-density lipoprotein; HDL, high-density lipoprotein; TG, triglyceride; HbA1c, hemoglobin A1c.

forest model, which calculates the individual DVT risk probability of the patient. Patients are stratified into high-risk (risk ≥20%) and low-risk categories based on a predefined threshold. All standardized baseline admission predictors (including demographic characteristics, comorbidities, diabetes-specific laboratory markers, routine laboratory parameters, admission functional status, fracture classification and Charlson Comorbidity Index) are preprocessed uniformly and input into the internally validated random forest ensemble model. The model generates a continuous individualized predicted

Table V. Predictive performance (AUC) of the random forest model across prespecified subgroups, with subgroup sample sizes and DVT event counts.

Subgroup	n	DVT events, n (%)	AUC (95% CI)
Overall	400	66 (16.5)	0.92 (0.87-0.96)
Diabetic	130	34 (26.2)	0.94 (0.89-0.98)
Non-diabetic	270	32 (11.9)	0.90 (0.85-0.94)
Age <80 years	220*	28 (12.7)[a]	0.91 (0.86-0.95)
Age ≥80 years	180*	38 (21.1)[a]	0.92 (0.87-0.96)
Men	180*	30 (16.7)[a]	0.91 (0.86-0.95)
Women	220*	36 (16.4)[a]	0.92 (0.87-0.96)
BMI <25 kg/m^2	280*	48 (17.1)[a]	0.93 (0.88-0.97)
BMI ≥25 kg/m^2	120*	18 (15.0)[a]	0.90 (0.85-0.94)
Independent mobility	150*	14 (9.3)[a]	0.88 (0.82-0.93)
Assisted/bedridden	250*	52 (20.8)[a]	0.93 (0.88-0.97)
HbA1c <7.0%	292	38 (13.0)	0.89 (0.84-0.94)
HbA1c ≥7.0%	108	28 (25.9)	0.95 (0.91-0.99)
LDL <3.0 mmol/l	240*	30 (12.5)[a]	0.90 (0.85-0.94)
LDL ≥3.0 mmol/l	160*	36 (22.5)[a]	0.93 (0.88-0.97)
Wait ≤3 days	250*	30 (12.0)[a]	0.89 (0.84-0.94)
Wait >3 days	150*	36 (24.0)[a]	0.93 (0.88-0.97)

[a]Estimates based on overall cohort proportions. For the HbA1c ≥7.0% subgroup (n=108; events=28), the AUC of 0.95 should be interpreted with caution due to the moderate sample size and determination in larger cohorts is warranted. AUC, area under curve; DVT, deep vein thrombosis; LDL, low-density lipoprotein; HbA1c, hemoglobin A1c.

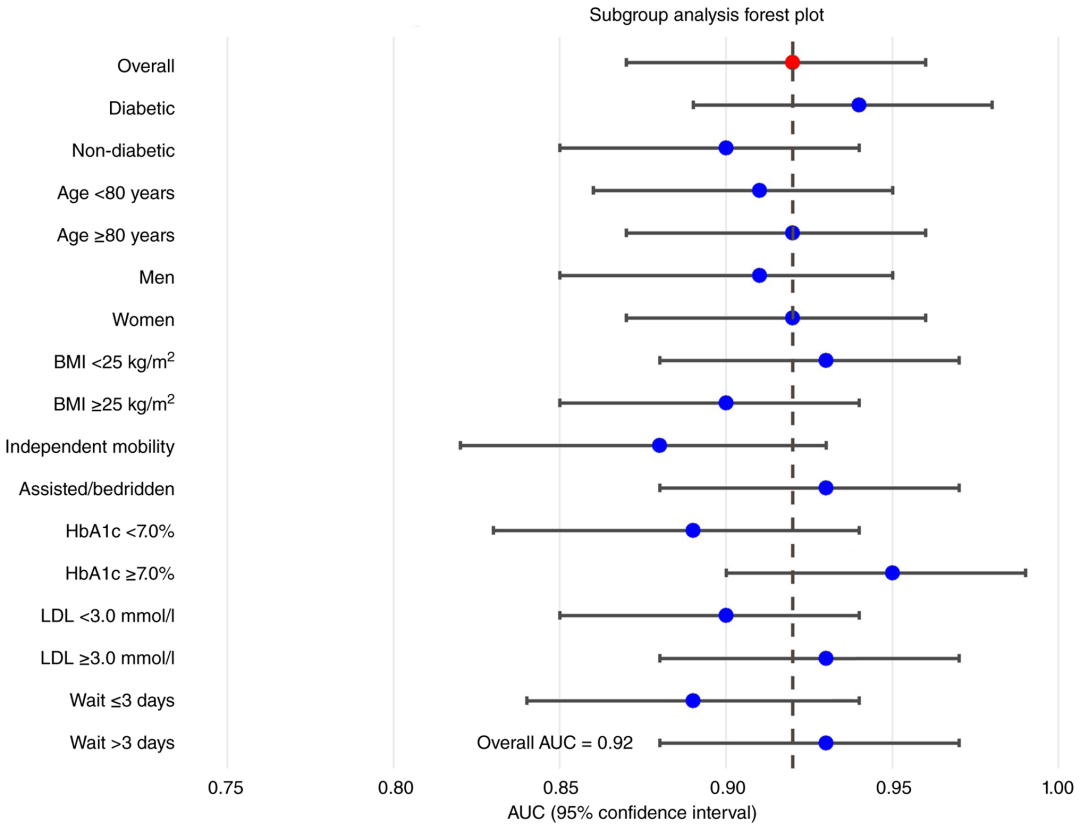

Figure 4. Subgroup analysis forest plot. Forest plot showing the AUC values of the random forest model across clinically relevant subgroups. The subgroups are: Overall, diabetic, non-diabetic, age <80 years, age ≥80 years, men, women, BMI <25 kg/m^2, BMI ≥25 kg/m^2, independent mobility, assisted/bedridden, HbA1c <7.0%, HbA1c ≥7.0%, LDL <3.0 mmol/l, LDL ≥3.0 mmol/l, wait ≤3 days and wait >3 days. The overall AUC (red dot) is 0.92; individual subgroup estimates are indicated by blue dots. The vertical dashed line represents the overall AUC. LDL, low-density lipoprotein; AUC, area under curve; HbA1c, hemoglobin A1c.

Table VI. Sensitivity analysis results.

Analysis type	Description	Random forest AUC (95% CI)
Primary analysis	Full dataset with MICE imputation (n=400)	0.92 (0.87-0.96)
Complete-case analysis	Patients with no missing data (n=352)	0.91 (0.86-0.95)
Random seed stability		
Seed 1	Random split 1	0.92 (0.87-0.96)
Seed 2	Random split 2	0.91 (0.86-0.95)
Seed 3	Random split 3	0.93 (0.88-0.97)
Seed 4	Random split 4	0.90 (0.85-0.94)
Seed 5	Random split 5	0.92 (0.87-0.96)
	Range	0.90-0.93
	Mean ± SD	0.92±0.01
Alternative imputation	Mean imputation instead of MICE	0.91 (0.86-0.95)
Univariate filtering	Model trained only on variables with P<0.05 from univariate analysis	0.91 (0.86-0.95)
Excluding lipid parameters	Model trained without lipid profile variables	0.88 (0.83-0.92)
Excluding diabetes variables	Model trained without diabetes-specific variables	0.85 (0.80-0.90)
10-fold cross-validation	Cross-validated AUC on training set	0.91 (0.87-0.95)

AUC, area under the curve; MICE, multiple imputation by chained equations.

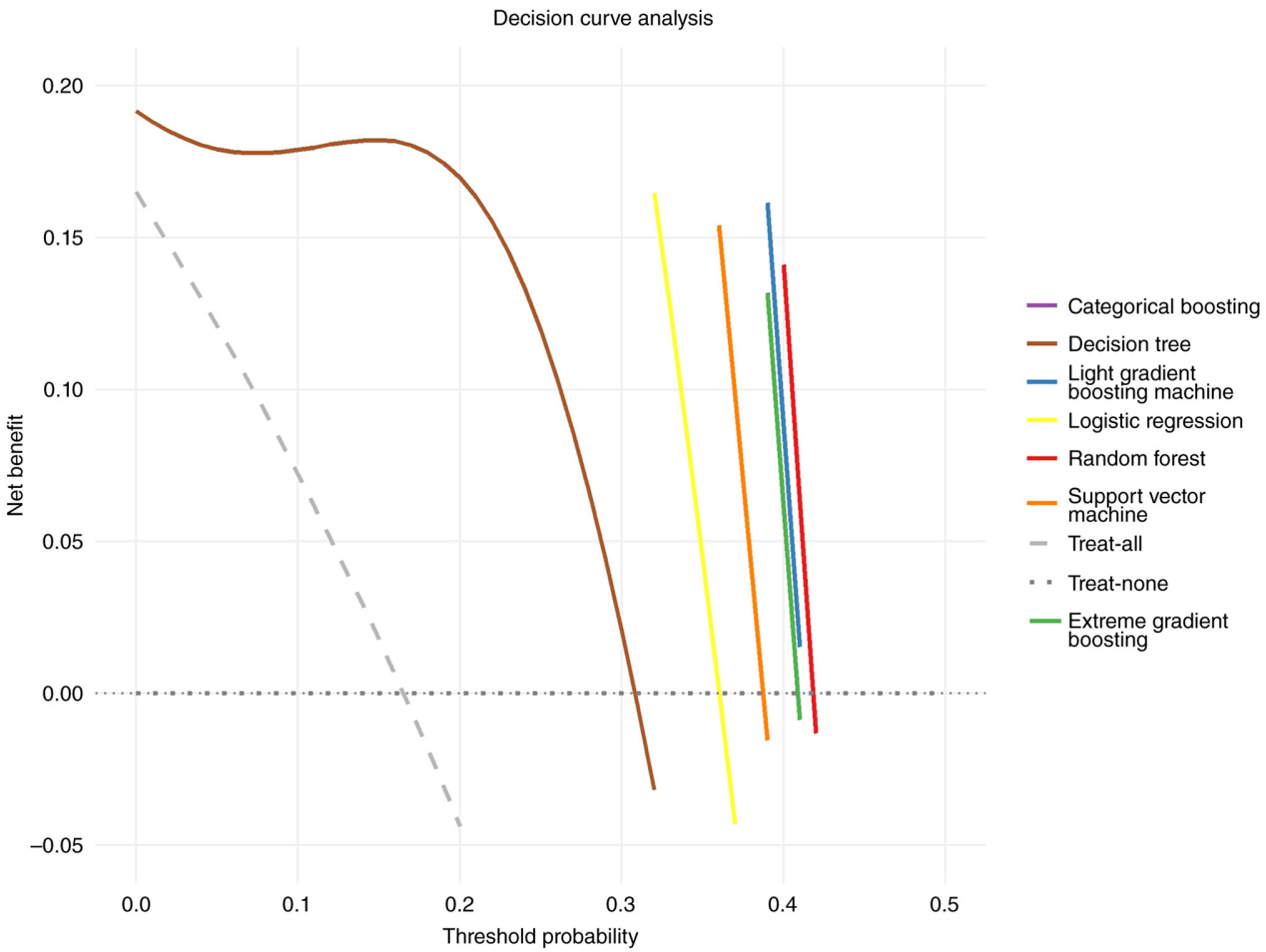

Figure 5. Decision curve analysis. Decision curves depicting the net benefit of seven machine learning models (CatBoost, Decision Tree, LightGBM, logistic regression, random forest, SVM and XGBoost) together with the treat-all and treat-none strategies across threshold probabilities from 0-0.5. The key identifies each model and strategy by color. SVM, support vector machine; XGBoost, Extreme Gradient Boosting.

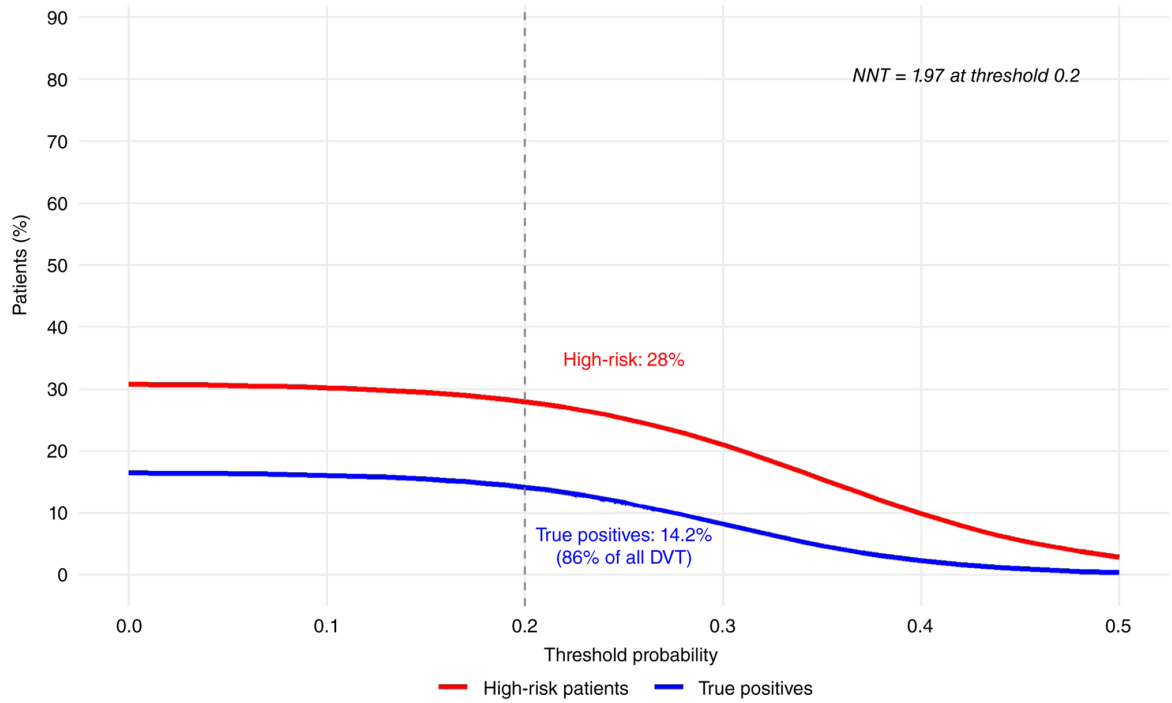

Figure 6. Clinical impact curve for the random forest model. The clinical impact curve illustrates the proportion of patients classified as high-risk (red line) and the proportion of true positives among them (blue line) across a range of risk thresholds. At a threshold of 0.2, ~28% of patients would be identified as high-riska nd 14.2% of all patients would be true positives, corresponding to a sensitivity of 86% (86% of all DVT cases are captured). The NNT at this threshold is 1.97, indicating that 1.97 patients need to be classified as high-risk and receive intervention to prevent one DVT event. The vertical dashed line marks the 0.2 threshold. DVT, deep vein thrombosis; NNT, needed to treat.

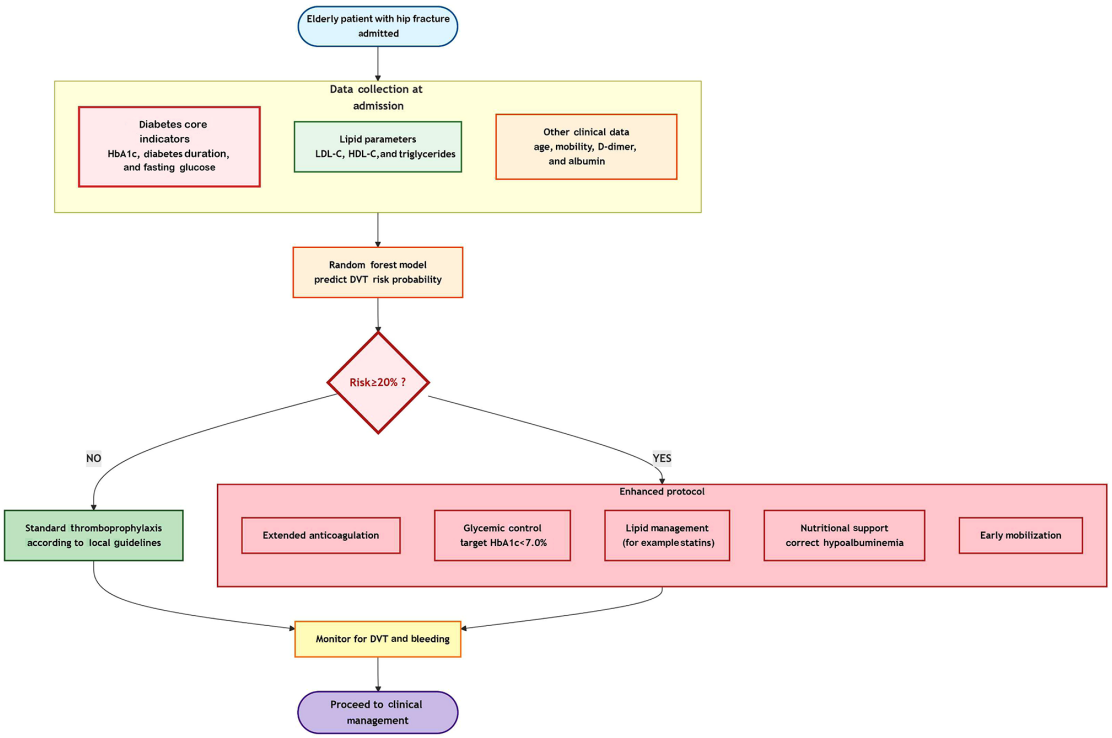

Figure 7. Proposed clinical decision pathway for DVT risk stratification and prevention at admission. This conceptual flowchart outlines a hypothesis-generating framework based on the present exploratory findings. External and prospective validation are required before any clinical implementation. The pathway is intended for research purposes to guide future studies. Data available on admission, including diabetes-specific variables (HbA1c, diabetes duration and fasting glucose) and lipid parameters (LDL cholesterol, HDL cholesterol and triglycerides), are collected and input into a validated random forest model to predict DVT risk. Patients are stratified into high-risk and low-risk categories based on a predefined threshold. High-risk patients trigger an enhanced thromboprophylaxis protocol encompassing extended anticoagulation, glycemic optimization targeting HbA1c <7.0%, lipid management, nutritional support to correct hypoalbuminemia and early mobilization protocols. Patients at low risk receive standard thromboprophylaxis according to local guidelines. This pathway enables early, personalized risk stratification and facilitates timely implementation of targeted preventive measures. LDL, low-density lipoprotein; HDL, high-density lipoprotein; DVT, deep vein thrombosis; HbA1c, hemoglobin A1c.

probability (range, 0-1) of postoperative DVT for each patient based on the full feature vector. Patients are stratified into two risk tiers using a predefined 20% probability threshold. This cutoff value was determined by maximizing Youden's J statistic on the internal validation ROC curve to balance screening sensitivity and specificity for clinical practice. Patients with predicted DVT risk ≥20% are categorized as high-risk candidates who require enhanced, intensified thromboprophylaxis, comprising strict glycemic control (target HbA1c <7%), extended anticoagulation, lipid management with statins, nutritional support to correct hypoalbuminemia and early mobilization. Patients with predicted risk <20% fall into the low-risk group eligible for standard routine venous thromboembolism prevention strategies, for example,with low-molecular-weight heparin during hospitalization and routine care. All patients are monitored for DVT and bleeding events throughout their hospital stay. This pathway enables early, personalized risk stratification and facilitates timely implementation of targeted preventive measures.

Discussion

A total of seven distinct ML models for predicting DVT in elderly hip fracture patients were successfully developed and internally validated in the present study using clinical data available at the time of admission. The random forest algorithm demonstrated improved predictive performance with an AUC of 0.92, significantly outperforming traditional logistic regression and demonstrating strong performance compared to other ensemble methods. The present comprehensive feature importance analysis, enhanced by SHAP interpretability methods, consistently identified diabetes-associated variables; HbA1c, diabetes duration and fasting glucose, as among the most important predictors, alongside D-dimer, age, albumin, time from admission to surgery and lipid parameters including LDL cholesterol and triglycerides. Diabetes-associated factors collectively accounted for ~30% of total model importance, while lipid parameters contributed an additional 18%, underscoring the notable role of metabolic dysregulation in DVT pathogenesis in this population. The model demonstrated robust performance across key subgroups, with particularly strong discrimination in diabetic patients (AUC of 0.94), those with HbA1c ≥7.0% (AUC of 0.95) and patients with impaired mobility (AUC of 0.93) and exhibited goof calibration and clinical utility. The findings of the present study aligned with and markedly extended previous research in DVT prediction among patients with hip fracture. The DVT incidence of 16.5% in the present cohort is consistent with large meta-analyses reporting rates of 16.6% (5,6), determining the representativeness of the present study population. The identification of diabetes-associated variables as dominant predictors corroborates growing evidence associating hyperglycemia and insulin resistance to hypercoagulability and thrombotic risk (7-10). The strong predictive value of HbA1c, a marker of chronic glycemic control, suggested that cumulative exposure to hyperglycemia may be more important compared with acute glucose levels on admission, consistent with studies demonstrating an association between HbA1c and cardiovascular events (9,12). The observed threshold effect at HbA1c 7.0% aligned with current diabetes management guidelines and suggested that achieving adequate glycemic control may have important benefits for thrombosis prevention beyond microvascular complications (13).

The consistent importance of nutritional status as reflected by albumin aligned with established literature on hypoalbuminemia as a marker of frailty and predictor of adverse outcomes in surgical patients (15). The present study found that hypoalbuminemia (OR=3.58) was among the strongest predictors reinforces the key role of nutritional optimization in perioperative care. The notable contribution of D-dimer and fibrinogen to prediction determined their well-established role as markers of thrombotic activation (14,17).

The marked contribution of lipid parameters to prediction extends previous research suggesting that dyslipidemia may influence thrombotic risk through effects on platelet function, coagulation factors and endothelial function (18,19). The present finding that LDL cholesterol and triglycerides were positively associated with DVT risk while HDL cholesterol demonstrated a protective effect, was consistent with the broader cardiovascular literature and suggested that the metabolic syndrome, encompassing both diabetes and dyslipidemia, created a particularly prothrombotic state (24).

Comparison with existing DVT risk assessment tools further highlights the incremental value of diabetes-associated features. Established tools, such as the Caprini score (33), assign only a single point for diabetes without accounting for glycemic control, disease duration, or treatment intensity. By contrast, the present model integrated HbA1c, diabetes duration, and insulin use, which collectively accounted for ~30% of total feature importance in the random forest model. Sensitivity analyses determined the unique contribution of these variables: Excluding them reduced the AUC from 0.92 to 0.85. Moreover, lipid parameters (LDL, HDL and triglycerides), which are absent from the majority of conventional DVT risk scores, contributed an additional 18% to model performance; their exclusion lowered the AUC to 0.88. These findings suggest that traditional risk assessment tools may systematically underestimate DVT risk in elderly patients with hip fracture with poorly controlled or long-standing diabetes and dyslipidemia.

The importance of process-associated factors, particularly time from admission to surgery, aligns with studies demonstrating that delayed surgical intervention increases complication risks in patients with hip fracture (25,34). The present finding, that each additional day of preoperative waiting increased DVT risk by 18%, provided strong impetus for systems-level interventions to minimize surgical delays.

The present study advances the field by systematically comparing numerous ML architectures and demonstrating the improved performance of ensemble methods compared with established statistical approaches (35,36). This performance advantage likely stems from the capacity of algorithms to capture complex nonlinear interactions among risk factors, such as the threshold effect of HbA1c, the interaction between diabetes duration and D-dimer and the synergistic effect of elevated LDL cholesterol and triglycerides, that may be missed by conventional models. The comparable performance of random forest, LightGBM, XGBoost and CatBoost suggested that modern gradient boosting methods all effectively capture

the underlying risk patterns, providing flexibility in model selection for future implementation.

In the context of hip fracture care pathways, the time of admission presents a key window for risk stratification and immediate initiation of thromboprophylaxis. The present study directly addressed this need by developing a pragmatic prediction tool that used data routinely available at admission and emphasized diabetes and lipid parameters as key modifiable risk factors (Fig. 7). Based on the present internal validation results, the random forest model demonstrated strong potential for clinical application in settings similar to those of the present study (tertiary general hospital orthopedic wards that admit large volumes of elderly patients with traumatic hip fracture, with standardized routine admission blood testing, including full lipid profiles and glycated hemoglobin assays, unified venous thrombosis screening protocols, and integrated multidisciplinary geriatric, endocrinology and orthopedic care systems. These institutions routinely implement standardized thromboprophylaxis protocols for hip fracture patients, and collect complete electronic medical record data covering demographics, comorbidities, functional status and perioperative outcomes, consistent with the single-center tertiary hospital environment of the present research cohort), as it utilizes commonly measured laboratory tests and clinical assessments without requiring specialized testing beyond standard admission panels, although this remains to be determined in external cohorts. Therefore, all implications discussed in the present study should be considered hypothesis-generating rather than practice-changing. The identification of HbA1c as the most important predictor has marked clinical implications. First, it suggests that preoperative glycemic optimization, targeting HbA1c <7.0%, could potentially reduce thrombotic risk, though this hypothesis requires prospective testing. For patients admitted with hip fracture, knowledge of HbA1c can immediately identify those with poorly controlled diabetes who warrant intensified monitoring and potentially more aggressive thromboprophylaxis. Second, the threshold effect at 7.0% provides a clear, actionable target for quality improvement initiatives. It should be acknowledged, however, that strict glycemic targets (HbA1c <7%) are often not achievable during the acute hip fracture admission due to the stress response, enteral access limitations and risk of hypoglycemia in frail elderly patients. Therefore, the proposed pathway should be viewed as a long-term risk-modification strategy and a research tool for future trials, rather than an immediate bedside protocol. To further translate SHAP findings into clinical action, the present study provides the following practical interpretations for key predictors. First, regarding HDL cholesterol, its protective effect means that low HDL levels (<1.0 mmol/l in men or <1.3 mmol/l in women) should alert clinicians to increased DVT risk. While raising HDL acutely during hospitalization is not feasible, this marker helps identify a high-risk metabolic phenotype that warrants intensified thromboprophylaxis and long-term cardiovascular risk management after recovery. Second, the HbA1c threshold of 7.0% serves as a clear warning sign: Patients with HbA1c ≥7.0% should trigger more aggressive glycemic control and enhanced DVT prevention. Third, albumin <35 g/l identifies patients with poor nutritional status who may benefit from nutritional support as part of thromboprophylaxis. Fourth, D-dimer levels >8,000 ng/ml show diminishing incremental risk, suggesting that the greatest predictive value lies in the moderately elevated range (3,000-8,000 ng/ml), where each unit increase confers notable risk elevation. Similarly, the strong contribution of lipid parameters suggests that patients with elevated LDL cholesterol and triglycerides or low HDL cholesterol represent a high-risk phenotype that may benefit from enhanced preventive strategies.

At a risk threshold of 20%, decision curve analysis indicates that the model would require classifying ~2 patients as high-risk to correctly identify one patient who will develop DVT (NNT=2.0), suggesting favorable cost-effectiveness for targeted thromboprophylaxis programs. The clinical impact curve further supports this, showing that at this threshold, 86% of actual DVT cases would be captured while avoiding unnecessary intervention in 72% of low-risk patients. This balance is particularly important in elderly patients, in whom the benefits of thromboprophylaxis must be carefully weighed against bleeding risks.

The identified key predictors provide clear targets for preemptive interventions. Diabetes optimization through intensified glycemic control during hospitalization may reduce thrombotic risk, potentially through enhanced fibrinolysis and reduced endothelial dysfunction (37). Lipid management, while typically a long-term intervention, may inform risk stratification and highlight patients who could benefit from statin therapy, which has been associated with reduced venous thromboembolism risk in certain studies (38,39). Nutritional support to correct hypoalbuminemia through appropriate supplementation could address the notable risk associated with poor nutritional status. Early mobilization protocols, implemented as soon as medically feasible, can minimize the prothrombotic effects of immobility (40). Systems-level interventions to reduce time from admission to surgery, potentially through dedicated hip fracture pathways, could directly address the process-associated risk we identified. For the highest-risk patients, particularly those with poorly controlled diabetes, adverse lipid profiles, hypoalbuminemia, elevated D-dimer or prolonged preoperative waiting, extended-duration anticoagulation post-discharge may be warranted (41).

The balanced sensitivity (0.85) and specificity (0.88) of the model support its utility for both identifying high-risk patients for intensive prevention strategies and avoiding unnecessary interventions in low-risk individuals, enabling efficient resource allocation in busy clinical settings. In the future, after external validation, integration with electronic health records might provide real-time risk scores at admission, potentially triggering alerts for high-risk patients and prompting evidence-based interventions through clinical decision support systems. However, such implementation is not yet supported by the current evidence and requires further study.

The present study exhibits a number of limitations that warrant careful consideration when interpreting the findings, as well as notable strengths that support its internal validity. The present study possesses certain strengths. First, the comprehensive assessment of variables available at admission, including detailed diabetes-specific measures (HbA1c, duration and treatment) and a full lipid profile, provide a more nuanced risk profile compared with previous hip fracture DVT prediction models that merely captured binary diabetes

presence/absence or isolated single lipid markers without complete lipid panels. Second, the inclusion of consecutive patients with strict eligibility criteria minimized selection bias and ensured a representative sample of the target population at our institution, thereby enhancing the internal validity of the present findings. Third, the rigorous comparison of seven distinct ML architectures with systematic hyperparameter tuning and 10-fold cross-validation ensures robust algorithm selection. Fourth, the application of SHAP analysis for model interpretability represents a notable advance, revealing nonlinear associations and interactions that would be undetectable with conventional methods. Fifth, extensive subgroup and sensitivity analyses determine the robustness and consistency of findings across clinically relevant populations and analytic decisions. Sixth, the use of data obtainable at admission enhances clinical utility by enabling immediate risk stratification upon presentation, allowing timely initiation of preventive measures. Seventh, the sample size of 400, with 66 DVT events, provided adequate events per variable for stable model development while maintaining feasibility for a single-center study. Although the number of DVT events (n=66) in the present single-center study as modest, the robustness of the model was supported by rigorous 10-fold cross-validation and a series of sensitivity analyses (including different random seeds and complete-case analysis), which consistently demonstrated stable performance (AUC range: 0.90-0.93). Therefore, despite the limited sample size, the internal validity of the present findings is well grounded.

A number of limitations warrant consideration. The most key limitation is the lack of external validation, followed by four additional methodological concerns. First, the single-center retrospective design may limit generalizability, and this lack of external validation prevents any direct clinical application of the model at this stage. Although the present cohort characteristics align with broader hip fracture populations and the DVT rate of 16.5% is consistent with published meta-analyses (42). Second, despite comprehensive variable collection, potential unmeasured confounders such as genetic thrombophilia, specific medication adherence, or subtle functional status measures not captured in electronic records may influence prediction accuracy. Third, the absence of a standardized prospective DVT screening protocol may introduce surveillance bias, though the inclusion of both screening and clinically indicated ultrasounds, combined with systematic chart review, mitigates this concern. Fourth, the sample size, while adequate for initial model development, may limit the detection of rare risk factors and the stability of estimates for some predictor interactions. Fifth, the present cohort included both surgically and conservatively managed patients and while this enhances generalizability, differences in management strategies could influence DVT risk in ways not fully captured by our predictors.

Notably, the absence of true external validation in this single-center study means the generalizability of the model to other settings is currently unknown and must be evaluated in independent cohorts. Therefore, the performance metrics reported in the present study should be interpreted as the capability of the model when applied to new patients from the same institution with characteristics similar to the development cohort. The present study deliberately avoid claims of broad generalizability and stress that external validation across different centers, populations and healthcare systems is a key next step to translate this predictive tool into clinical practice. Addressing these limitations through the future research directions outlined below is key in clinical translation.

Based on the present findings, a number of promising research directions emerge. First and foremost, external validation in diverse healthcare settings and populations is important in establishing generalizability and assess transportability. Multi-center collaborative studies should aim to evaluate model performance across institutions with different patient demographics, clinical practices and DVT rates. Second, prospective validation studies with standardized DVT screening protocols would provide the highest level of evidence and eliminate potential surveillance bias.

Third, the development of customized intervention protocols targeting the identified high-impact predictors represent a logical next step. Clinical trials could randomize high-risk patients identified by the model to intensified thromboprophylaxis vs. standard care, evaluating not only DVT reduction but also bleeding complications and functional outcomes. Fourth, implementation science studies should evaluate the integration of this tool into electronic health record systems for real-time risk assessment and clinical decision support at the point of admission, measuring impact on process metrics such as time to prophylaxis initiation, adherence to anticoagulation protocols and resource allocation.

Fifth, the combination of clinical variables with novel biomarkers may further enhance prediction accuracy. Thrombin generation assays, plasminogen activator inhibitor-1, inflammatory cytokines and genetic polymorphisms associated with thrombophilia could provide additional mechanistic insights and improve risk stratification. Sixth, prospective studies examining the association between glycemic and lipid optimization during hospitalization and DVT risk reduction would provide key evidence for causal inference and intervention design. Seventh, longitudinal studies examining the long-term outcomes of patients identified as high-risk, including post-thrombotic syndrome, recurrent venous thromboembolism and mortality, would provide valuable evidence for optimizing comprehensive care pathways. Eighth, health economic analyses should evaluate the cost-effectiveness of model-guided thromboprophylaxis strategies, considering both the costs of intervention and the savings from prevented complications. These future studies will directly address the current limitations of single-center design, lack of external validation and absence of prospective screening protocols.

Finally, the development of simplified risk scores based on the most important predictors identified through ML could facilitate bedside application in settings without access to electronic health records or advanced computational infrastructure. Such scores would require careful calibration and validation but could extend the benefits of risk stratification to a broader range of clinical settings.

In conclusion, the present findings should be viewed as hypothesis-generating and require determination in independent cohorts before any clinical translation. While the random forest model showed promise for DVT risk stratification in elderly hip fracture patients, external validation is important prior to implementation.

Acknowledgements

Not applicable.

Funding

The present study was supported by grants from the Tianjin Hospital Science and Technology Fund (grant no. Stjyyq2512) and the Tianjin Municipal Health Commission (grant nos: ZC20191 and TJWJ2024RC011).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

QS and PT designed the present study. QS, PS, PT, AC, YZ and QD performed the data collection and analyzed the data. QS and PT confirm the authenticity of all the raw data. QS drafted the manuscript. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

The present study was approved by the Institutional Review Board of Tianjin Hospital, China prior to data collection and analysis (Tianjin, China; approval no. 025-195). The requirement for individual informed consent was waived due to the retrospective nature of the study, which involved minimal risk to participants and used previously collected anonymized data. This decision was consistent with national regulations and institutional guidelines for retrospective chart review studies.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Zrour S, Belhaj Salem S, Ben Chekaya N, Grassa R, Bejia I and Abid A: Survival rate after osteoporotic proximal femur fractures. Tunis Med 102: 1078-1083, 2024 (In French).
2. Andaloro S, Cacciatore S, Risoli A, Comodo RM, Brancaccio V, Calvani R, Giusti S, Schlögl M, D'Angelo E, Tosato M, *et al*: Hip fracture as a systemic disease in older adults: A narrative review on multisystem implications and management. Med Sci (Basel) 13: 89, 2025.
3. Fa-Binefa M, Clara A, Lamas C and Elosua R: Mediterranean diet and risk of hip fracture: A systematic review and dose-response meta-analysis. Nutr Rev 83: 1133-1143, 2025.
4. Soro-García P and González-Gálvez N: Effects of progressive resistance training after hip fracture: A systematic review. J Funct Morphol Kinesiol 10: 54, 2025.
5. Geerts WH, Bergqvist D, Pineo GF, Heit JA, Samama CM, Lassen MR and Colwell CW: Prevention of venous thromboembolism: American college of chest physicians evidence-based clinical practice guidelines (8th edition). Chest 133 (Suppl 6): 381S-453S, 2008.
6. Zöller B, Li X, Sundquist J and Sundquist K: Risk of pulmonary embolism in patients with autoimmune disorders: A nationwide follow-up study from Sweden. Lancet 379: 244-249, 2012.
7. Anderson FA Jr and Spencer FA: Risk factors for venous thromboembolism. Circulation 107 (Suppl 1): I9-I16, 2003.
8. Zhang L, He M, Jia W, Xie W, Song Y, Wang H, Peng J, Li Y, Wang Z and Lin Z: Analysis of high-risk factors for preoperative DVT in elderly patients with simple hip fractures and construction of a nomogram prediction model. BMC Musculoskelet Disord 23: 441, 2022.
9. Piazza G, Goldhaber SZ, Kroll A, Goldberg RJ, Emery C and Spencer FA: Venous thromboembolism in patients with diabetes mellitus. Am J Med 125: 709-716, 2012.
10. Ageno W, Becattini C, Brighton T, Selby R and Kamphuisen PW: Cardiovascular risk factors and venous thromboembolism: A meta-analysis. Circulation 117: 93-102, 2008.
11. Yang CS and Tan Z: Construction and validation of a predictive model for preoperative lower extremity deep vein thrombosis risk in elderly hip fracture patients: An observational study. Medicine (Baltimore) 103: e39825, 2024.
12. Caprini JA: Risk assessment as a guide for thrombosis prophylaxis. Curr Opin Pulm Med 16: 448-452, 2010.
13. Doggen CJM, Smith NL, Lemaitre RN, Heckbert SR, Rosendaal FR and Psaty BM: Serum lipid levels and the risk of venous thrombosis. Arterioscler Thromb Vasc Biol 24: 1970-1975, 2004.
14. Moellmann HL, Alhammadi E, Boulghoudan S, Kuhlmann J, Mevissen A, Olbrich P, Rahm L and Frohnhofen H: Risk of sarcopenia, frailty and malnutrition as predictors of postoperative delirium in surgery. BMC Geriatr 24: 971, 2024.
15. Niu Y, Wang Q, Lu J, He P and Guo HT: Risk factors for postoperative delirium in orthopedic surgery patients: A systematic review and meta-analysis. Ann Med 57: 2534520, 2025.
16. Ge X, Yao L, Liu Y, Wang Y and Zhang F: Comparing machine learning models for predicting preoperative DVT incidence in elderly hypertensive patients with hip fractures: A retrospective analysis. Sci Rep 15: 13206, 2025.
17. Chen H, Yu D, Zhang J and Li J: Machine learning for prediction of postoperative delirium in adult patients: A systematic review and meta-analysis. Clin Ther 46: 1069-1081, 2024.
18. Yuan J, Zeng Q, Li J, Cong Z and Zhang Y: Machine learning applications in sports injury prediction: A narrative review. Sci Prog 108: 368504251385956, 2025.
19. Rozera T, Pasolli E, Segata N and Ianiro G: Machine learning and artificial intelligence in the multi-omics approach to gut microbiota. Gastroenterology 169: 487-501, 2025.
20. Ramos MV: Reviewing the context of molecular modeling to enhance the application of machine learning technologies for safer bioinformatics. Protein Pept Lett 32: 772-775, 2025.
21. Dastan D, Soleymanekhtiari S and Ebadi A: Peptidic compound as DNA binding agent: In silico fragment-based design, machine learning, molecular modeling, synthesis, and DNA binding Evaluation. Protein Pept Lett 31: 332-344, 2024.
22. Gui C, Gao Y, Zhang R and Zhou G: Bioinformatics analysis of lactylation-related biomarkers and potential pathogenesis mechanisms in age-related macular degeneration. Curr Genomics 26: 191-209, 2025.
23. Gomase VS, Dhamane SP, Kemkar KR, Kakade PG and Sakhare AD: Immunoproteomics: Approach to diagnostic and vaccine development. Protein Pept Lett 31: 773-795, 2024.
24. Mann J, Lyons M, O'Rourke J and Davies S: Machine learning or traditional statistical methods for predictive modelling in perioperative medicine: A narrative review. J Clin Anesth 102: 111782, 2025.
25. Holler E, Ludema C, Ben Miled Z, Rosenberg M, Kalbaugh C, Boustani M and Mohanty S: Development and validation of a routine electronic health record-based delirium prediction model for surgical patients without dementia: Retrospective case-control study. JMIR Perioper Med 8: e59422, 2025.
26. Needleman L, Cronan JJ, Lilly MP, Merli GJ, Adhikari S, Hertzberg BS, DeJong MR, Streiff MB and Meissner MH: Ultrasound for lower extremity deep venous thrombosis: Multidisciplinary recommendations from the society of radiologists in ultrasound consensus conference. Circulation 137: 1505-1515, 2018.
27. Tamariz L, Harkins T and Nair V: A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. Pharmacoepidemiol Drug Saf 21 (Suppl 1): S154-S162, 2012.
28. Drosdowsky A and Gough K: The charlson comorbidity index: Problems with use in epidemiological research. J Clin Epidemiol 148: 174-177, 2022.

29. Liu C, Peng XX, Cai SY, Liu YL, Zhang C and Hu F: Development of a pre-processing workflow for real world data derived from multicenter clinical laboratories. Zhonghua Liu Xing Bing Xue Za Zhi 46: 296-306, 2025 (In Chinese).
30. Namjoo-Moghadam A, Abedi V, Avula V, Ashjazadeh N, Hooshmandi E, Abedinpour N, Rahimian Z, Borhani-Haghighi A and Zand R: Machine learning-based cerebral venous thrombosis diagnosis with clinical data. J Stroke Cerebrovasc Dis 33: 107848, 2024.
31. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, Van Smeden M, *et al*: TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385: e078378, 2024.
32. Collins GS, Reitsma JB, Altman DG and Moons KG: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. BMJ 350: g7594, 2015.
33. Hayssen H, Cires-Drouet R, Englum B, Nguyen P, Sahoo S, Mayorga-Carlin M, Siddiqui T, Turner D, Yesha Y, Sorkin JD and Lal BK: Systematic review of venous thromboembolism risk categories derived from Caprini score. J Vasc Surg Venous Lymphat Disord 10: 1401-1409.e7, 2022.
34. Clinkenbeard K, Bossle K, Pape T, Woltenberg LN and Saha S: Time to hip fracture surgery and mortality. South Med J 116: 274-278, 2023.
35. Ten Cate V, Prochaska JH, Schulz A, Nagler M, Robles AP, Jurk K, Koeck T, Rapp S, Düber C, Münzel T, *et al*: Clinical profile and outcome of isolated pulmonary embolism: A systematic review and meta-analysis. EClinicalMedicine 59: 101973, 2023.
36. Ma R, Yu W, Tian J, Tang Y, Fang H, Ming X and Liu H: Machine learning in the prediction of venous thromboembolism: Systematic review and meta-analysis. J Med Internet Res 27: e77339, 2025.
37. Selvin E, Steffes MW, Zhu H, Matsushita K, Wagenknecht L, Pankow J, Coresh J and Brancati FL: Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. N Engl J Med 362: 800-811, 2010.
38. Davis JW, Weller SC, Porterfield L, Chen L and Wilkinson GS: Statin use and the risk of venous thromboembolism in women taking hormone therapy. JAMA Netw Open 6: e2348213, 2023.
39. Pencina KM, Thanassoulis G, Pencina MJ, Toth PP and Sniderman AD: Hemoglobin A1c and abdominal obesity as predictors of diabetes and ASCVD in individuals with prediabetes in UK Biobank: A prospective observational study. Cardiovasc Diabetol 23: 448, 2024.
40. Jiao X, Zhang Q, Peng P and Shen Y: HbA1c is a predictive factor of severe coronary stenosis and major adverse cardiovascular events in patients with both type 2 diabetes and coronary heart disease. Diabetol Metab Syndr 15: 50, 2023.
41. Merrell LA, Esper GW, Ganta A, Egol KA and Konda SR: Impact of poorly controlled diabetes and glycosylated hemoglobin values in geriatric hip fracture mortality risk assessment. Cureus 15: e36422, 2023.
42. Kanchanabat B, Stapanavatr W, Meknavin S, Soorapanth C, Sumanasrethakul C and Kanchanasuttirak P: Systematic review and meta-analysis on the rate of postoperative venous thromboembolism in orthopaedic surgery in Asian patients without thromboprophylaxis. Br J Surg 98: 1356-1364, 2011.