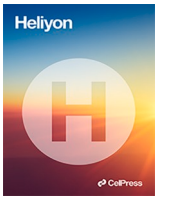
Heliyon 10 (2024) e29517

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Prediction model of deep vein thrombosis risk after lower extremity orthopedic surgery

Jiannan Zhang [a], Yang Shao [a], Hongmei Zhou [a], Ronghua Li [a], Jie Xu [b,c], Zhongzhou Xiao [b], Lu Lu [b], Liangyu Cai [a,*]

[a] Department of Anesthesiology, Wuxi TCM Hospital, Wuxi, 214071, PR China
[b] Shanghai Artificial Intelligence Laboratory, Shanghai, 200030, PR China
[c] Université de Montpellier, Montpellier, Languedoc-Roussillon, France

ARTICLE INFO

Keywords:
Lower extremity orthopedic surgery
Deep venous thrombosis
Risk prediction
D-dimer

ABSTRACT

Purpose: This investigation was conceived to engineer and appraise a pioneering clinical nomogram, crafted to bridge the extant chasm in literature regarding the postoperative risk stratification for deep vein thrombosis (DVT) in the aftermath of lower extremity orthopedic procedures. This novel tool offers a sophisticated and discerning algorithm for risk prediction, heretofore unmet by existing methodologies.

Methods: In this retrospective observational study, clinical records of hospitalized patients who underwent lower extremity orthopedic surgery were collected at the Wuxi TCM Hospital Affiliated to the Nanjing University of Chinese Medicine between Jan 2017 and Oct 2019. The univariate and multivariate analysis with the backward stepwise method was applied to select features for the predictive nomogram. The performance of the nomogram was evaluated with respect to its discriminant capability, calibration ability, and clinical utility.

Result: A total of 5773 in-hospital patients were eligible for the study, with the incidence of deep vein thrombosis being approximately 1 % in this population. Among 31 variables included, 5 of them were identified to be the predictive features in the nomogram, including age, mean corpuscular hemoglobin concentration (MCHC), D-dimer, platelet distribution width (PDW), and thrombin time (TT). The area under the receiver operating characteristic (ROC) curve in the training and validation cohort was 85.9 % (95%CI: 79.96 %–90.04 %) and 85.7 % (95%CI: 78.96 %–90.69 %), respectively. Both the calibration curves and decision curve analysis demonstrated the overall satisfactory performance of the model.

Conclusion: Our groundbreaking nomogram is distinguished by its unparalleled accuracy in discriminative and calibrating functions, complemented by its tangible clinical applicability. This innovative instrument is set to empower clinicians with a robust framework for the accurate forecasting of postoperative DVT, thus facilitating the crafting of bespoke and prompt therapeutic strategies, aligning with the rigorous standards upheld by the most esteemed biomedical journals.

* Corresponding author. Department of Anesthesiology, Wuxi TCM Hospital, No. 8, Zhongnan West Road, Wuxi, 214071, Jiangsu, PR China.
E-mail addresses: 584363453@qq.com (J. Zhang), 375803522@qq.com (Y. Shao), 4094595@qq.com (H. Zhou), lironghua1990@163.com (R. Li), xiaochongzhou@pjlab.org.cn (J. Xu), xujie@pjlab.org.cn (Z. Xiao), lulu@pjlab.org.cn (L. Lu), wxzy018@njucm.edu.cn (L. Cai).

https://doi.org/10.1016/j.heliyon.2024.e29517
Received 7 October 2023; Received in revised form 8 April 2024; Accepted 9 April 2024
Available online 16 April 2024

1. Introduction

An expanding demand for lower extremity orthopedic surgery (LEOS), including elective total hip replacement (THR), total knee replacement (TKR) or hip fracture surgery (HFS) is reported mostly because of the increasing overall life expectancy as well as bone and joint diseases among older people such as rheumatoid arthritis [1–3]. In recent years in China, exercise culture making the orthopedic surgery climbing progressively as an inevitable part of middle-old life [4]. Considering the basic incidence rate that nearly one million adults undergo LEOS each year in the United States [5], it is understandable to discover the corresponding high hospitalizations up to 0.6 million due to deep vein thrombosis (DVT) and pulmonary embolism (PE) [6], and asymptomatic DVT incidence of 36%–84 % after LEOS without pharmacological thromboprophylaxis [7]. However, different from those in western countries, only 17%–43 % incidence of asymptomatic DVT could be observed in Asia patients [8], while the incidence of registered DVT after LEOS was under 0.3 % in Korea [9]. A recent prospective multicenter study indicated that the incidence of hospital DVT was 0.9 % in China [10]. The worldwide consensus was to prevent the critical morbidity and mortality in patients undergoing LEOS, postoperative venous thromboembolism (VTE), which more likely progressed to DVT [11,12], received significant attention [7,13–15]. DVT, occurring of the perioperative period of orthopedic surgery, poses a number of burdens, including prolongation of hospital stay, affecting patients' recovery process, and potentially leading to fatal PE [16]. It is reported that no matter how the lower extremity was injured or not, DVT could occur [17]. The higher incidences of DVT were 54%–62 % [18,19] in the injured patients compared with that of 9.63 % preoperatively to 20.29 % postoperatively in the uninjured ones [20,21], while pooled calculated of either the uninjured or injured lower extremity as 29.4 % preoperatively to 32.8 % postoperatively [22]. This partly reflected the estimated incidences of DVT for those patients undergoing surgeries after injury.

DVT, which occurs often in the lower extremity, and PE constitute VTE. It is usually followed by two associated diseases of chronic thromboembolic pulmonary hypertension [23] and mild to moderate postthrombotic syndrome (PTS) [24]. As DVT passes through the lungs, its progression can complicate treatment, resulting in life-threatening cardiac collapse [25]. Given the high incidence of this disease after LEOS and the serious adverse outcomes described above, diagnosis and even prediction of DVT becomes crucial during the treatment period. Based on the decades of development of diagnosis-related radiology, growing clinical experience, and artificial intelligence (AI) algorithms, a large number of studies have proved that imaging tests, clinical prediction rules, D-dimer testing, and algorithm could be efficient for DVT diagnosis to some extent [26]. Proximal compression ultrasonography (US) better than over whole-leg US is the imaging test choice [27], which can be more useful with the addition of Doppler, but serial testing was reported not cost-effective [28,29]. Another diagnosis method of DVT, magnetic resonance direct thrombus imaging, is a noninvasive test but only accurate below the knee [30]. Clinically, multitudinous scoring systems have been designed or validated to estimate the pre-test probability of DVT, such as the Wells criteria [31–33]. With a higher sensitivity and relatively lower specificity, the D-dimer assay is usually used to diagnose DVT when ruling out low-risk patients [31,34].

Taken together, the combination of pretest probability assessment, D-dimer, and US was recommended as the first-line strategy for the diagnosis of DVT since 2012 [35]. Although anticoagulant therapy could decrease the risk of recurrent thrombosis, it also increases

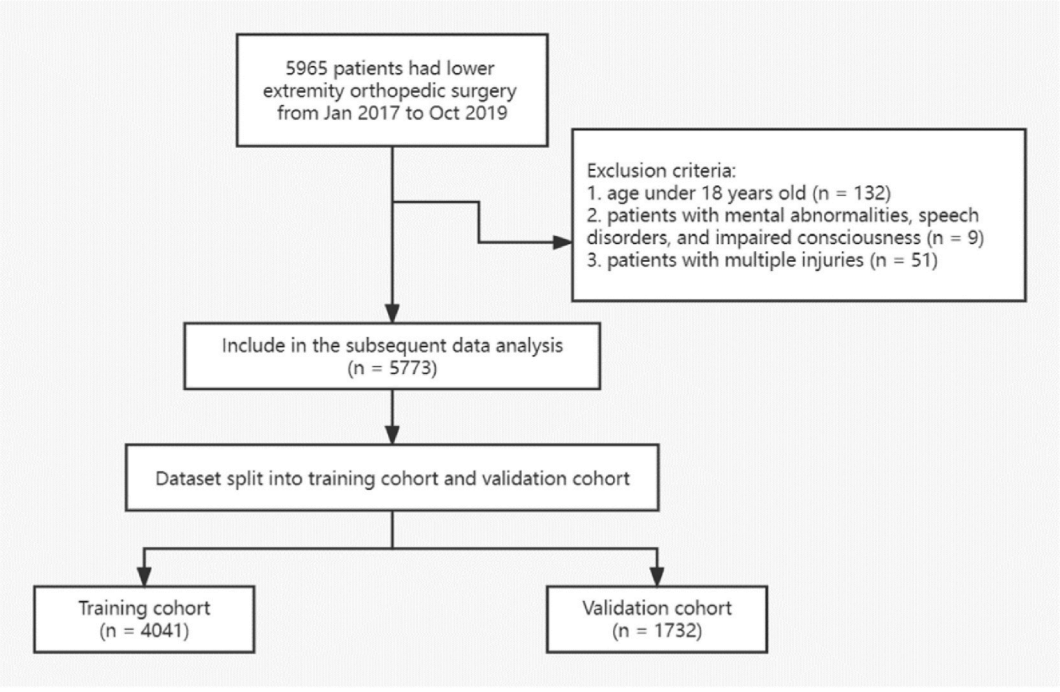


Fig. 1. The workflow of study population.

the risk of major hemorrhage, which makes the DVT prediction of more significances [26]. The construction of a prediction model for DVT risk has been continuously evolving in this decade. Among these approaches to DVT diagnosis, we found age and gender were both key features [36–39]. However, most of the current machine learning approaches were only focused on knee arthroscopy [40], THR or TKR [41–44], or lower extremity fractures like distal leg fractures and spinal fractures [45,46]. There are few studies on the prediction model of DVT risk after LEOS, possibly because of difficulties with data collection or uncertain effect on composite endpoint diagnosis. To address this potential issue, the missing values was integrated through the multivariate imputation by chained equations (MICE). The purpose of our study was to identify potential predictors for DVT after LEOS, which would fill the gap of prediction model construction of DVT risk in the perioperative period of LEOS for the first time.

2. Materials and methods

The retrospective study was performed in accordance with the relevant guidelines and regulations with the principles of the Helsinki Declaration. The experiments were approved by the Ethic Review Committee of Wuxi TCM Hospital Affiliated to the Nanjing University of Chinese Medicine (SWJW2019082907), and confirmed that patient informed consent was waived.

3. Study population

In this retrospective cohort study, we derived data between January 1, 2017 and October 31, 2019 from the inpatient medical record system at the Wuxi TCM Hospital Affiliated to the Nanjing University of Chinese Medicine. The database is manually maintained that collected data on demographic, lifestyle, laboratory tests, and intraoperative anesthesia of hospitalized patients. We included patients who had undergone lower extremity orthopedic surgery following the guideline of ICD-10-CM CODES. The exclusion criteria were as follows: (1) age under 18 years old; (2) patients with mental abnormalities, speech disorders, and impaired consciousness; (3) patients with multiple injuries. The workflow is shown in Fig. 1. In terms of the sample size, based on the requirement for developing a clinical prediction model [47], our sample size is surely sufficient.

4. Variable selection

We searched for variables contributing to DVT that were reported in published articles or reviews without time and language restrictions, and can be examined routinely in different settings with various clinical experiences. In addition, to enhance practicality and interpretability, we decided to choose reasonable and clinically relevant variables based on the Delphi method. In this case, we collected data on demographic information, chronic comorbidities, and laboratory biomarkers.

We collected data on patients' demographics including age, operation time, and the calculated body mass index (BMI) using height and weight. Operation time as one of the continuous variables, was transformed into discrete variables using pre-specified cut-offs point ($\leq$2h; 2–3h; $\geq$3h). The chronic comorbidities included smoking, alcohol drinking, a history of blood transfusion, hypertension, and diabetes. The laboratory biomarkers included red blood cell (RBC), white blood cell (WBC), platelet (PLT), hemoglobin (HGB), lymphocyte (LYM), neutrophil (NEU), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), platelet distribution width (PDW), red blood cell distribution width (RDW), glutamic pyruvic transaminase (ALT), glutamic oxaloacetic transaminase (AST), uric acid (UA), cholinesterase (CHE), total protein (TP), Globulin, D-dimer, Fibrinogen (FIB), activated partial thromboplastin time (APTT), thrombin time (TT), and Prothrombin time (PT). The latest data was selected if multiple medical records were presented.

5. Outcomes

The primary outcome of this study was the incidence of lower extremity deep venous thrombosis (LDVT) after LEOS, which defined according the guideline [48] as: (1) edema or sharp pain in affected limbs with significant tenderness in the femoral triangle or lower legs; the skin is dark red with increased temperature in affected limbs; irritability of superficial veins; positive Homans sign; (2) history of bed rest, surgery, trauma, malignancy, travel, thrombophilia, VTE, pregnancy and other risk factors for DVT; (3) the diagnosis can be confirmed by ultrasound Doppler and venogram; (4) acute plasma D-dimer is higher than normal. Acute arterial embolism, acute lymphangitis, erysipelas, and primary pelvic tumors, surgical hematoma or myofibrositis of the lower leg are excluded.

6. Data processing

The data were extracted and stored in the electronic spreadsheet, and were exported to R software version 4.2.1 for preliminary processing. Before the inclusion and exclusion, unstructured and abnormal data were screened and handled manually, by either dropping or replacing appropriately. Afterward, a slight proportion of missing was found on the variables age (1 %) and BMI (12 %), and was imputed by the mean values on the basis of data distribution.

Other missing values were found on laboratory biomarkers as well with varied missing proportions, which were summarized in the Supplemental Materials (Table S1). Assuming the missing was missing at random (MAR), we imputed the missing values through the MICE.

7. Statistical analysis

All statistical analysis was performed using the R software version 4.2.1 (2022-06-23 ucrt). The normality of the variables was accessed using the Kolmogorov-Smirnov test. Continuous variables were presented as mean (standard deviation, SD) if normal assumption satisfied, otherwise as median (and interquartile ranges, IQR). Counts and percentages were used for categorical variables. Student t-test or Wilcoxon rank sum test was used for continuous variables, and Chi-squared or Fisher's exact test was used for categorical variables, as appropriate.

Data were first split randomly into a training cohort and validation cohort according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement. Covariates with $p < 0.05$ in the univariate analysis remained in the multivariate logistics regression analysis. A backward stepwise selection was then applied to identify the potential predictor for constructing the nomogram. Each predictive factor retained in the nomogram was accessed independently as well by plotting the ROC curves. To evaluate the discriminant capability of the nomogram, the conformance index (C-index), and the area under the receiver operating characteristic (ROC) curve (AUC) were measured and validated both internally and externally with bootstrapping. In addition, the calibration curve was employed to access the calibration ability of the nomogram, incorporating the Hosmer and Lemeshow test. The clinical usefulness of the nomogram was explored by the decision curve analysis (DCA), which compares the threshold probabilities and the net benefit. $P < 0.05$ was considered statistically significant.

A sensitivity analysis was conducted to verify the validity of the MAR assumption, with multivariate logistics regression analysis performed on the original dataset.

8. Results

8.1. Characteristics of patients

In this study, a total of 5773 clinical records with lower extremity orthopedic surgery were selected on the basis of the exclusion criteria, of which 4041 were included in the training cohort and 1732 in the validation cohort (Fig. 1). There are no statistically significant differences between the two cohorts in DVT incidence (0.74 % vs. 1.27 %, $p = 0.07$).

Among the training dataset, the median age of patients with or without DVT was 60 (IQR, 48–70) and 75 (67–82), respectively. 837 (20.7 %) patients were found to have hypertension and 303 (7.5 %) had diabetes. With similar distribution, the two datasets showed alike behaviors with identical significance of variables except for Hypertension. Details on patients' baseline characteristics are summarized in Table 1.

Clinical characteristics based on risk groups are shown in Table 2. Within the training dataset, the values of variables including RBC, HGB, LYM, HCT, MCHC, PDW, UA, and TT were found to be significantly lower in the non-DVT groups compared to the DVT risk group, while WBC, NEU, D-dimer, and FIB were deemed to be statistically significantly higher. Similarly, the validation dataset provided comparable results even though there were a few inconsistencies in variables significance including WBC, RDW, CHE, and TT.

Table 1
Baseline characteristics of patients.

	Training Cohort			Validation Cohort		
Characteristic	DVT[a] (n = 30)	Non-Dvt[a] (n = 4011)	p[b]	DVT[a] (n = 22)	Non-Dvt[a] (n = 1710)	p[b]
BMI	24.1 (22.2, 26.0)	24.0 (22.0, 25.4)	0.3	24.0 (23.4, 24.4)	24.0 (22.0, 25.6)	>0.9
Age	75 (67, 82)	60 (48, 70)	<0.001	71 (66, 68)	60 (49, 70)	<0.001
Smoking (Yes)	0 (0 %)	111 (2.8 %)	>0.9	1 (4.5 %)	56 (3.3 %)	0.5
Alcohol (Yes)	1 (3.3 %)	134 (3.3 %)	>0.9	0 (0 %)	62 (3.6 %)	>0.9
Blood Transfusion			>0.9			0.3
Yes	0 (0 %)	39 (1.0 %)		1 (4.5 %)	19 (1.1 %)	
No	29 (97 %)	3770 (94 %)		20 (91 %)	1607 (94 %)	
Unknown	1 (3.3 %)	202 (5.0 %)		1 (4.5 %)	84 (4.9 %)	
Hypertension			0.002			0.5
Yes	13 (43 %)	824 (21 %)		4 (18 %)	415 (24 %)	
No	17 (57 %)	3187 (79 %)		18 (82 %)	1295 (76 %)	
Diabetes			0.069			>0.9
Yes	5 (17 %)	298 (7.4 %)		1 (4.5 %)	138 (8.1 %)	
No	25 (58 %)	3713 (93 %)		21 (95 %)	1572 (92 %)	
Operation time			0.5			0.8
$< 2h$	17 (57 %)	2397 (60 %)		12 (55 %)	988 (58 %)	
$2-3h$	8 (27 %)	1205 (30 %)		7 (32 %)	558 (33 %)	
$\geq 3h$	5 (17 %)	409 (10 %)		3 (14 %)	164 (9.6 %)	

[a] Median (IQR); n (%).
[b] Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test.

Table 2
Clinical characteristics of patients.

	Training Cohort			Validation Cohort		
Characteristic	DVT[a] (n = 30)	Non-Dvt[a] (n = 4011)	p[b]	DVT[a] (n = 22)	Non-Dvt[a] (n = 1710)	p[b]
RBC	3.64 (0.60)	4.15 (0.67)	<0.001	3.52 (0.45)	4.14 (0.66)	<0.001
WBC	8.38 (6.90, 9.36)	7.38 (5.99, 9.07)	0.029	7.47 (6.12, 8.22)	7.35 (5.99, 8.79)	>0.9
PLT	201 (177, 283)	204 (167, 244)	0.3	216 (189, 274)	208 (170, 247)	0.5
HGB	108 (17)	124 (20)	<0.001	102 (13)	124 (20)	<0.001
LYM	15 (13, 18)	20 (15, 27)	<0.001	18 (15, 19)	21 (15, 28)	0.023
NEU	76 (73, 80)	72 (64, 77)	<0.001	74 (72, 77)	71 (64, 77)	0.038
HCT	33.3 (4.8)	37.5 (5.6)	<0.001	31.2 (3.7)	37.4 (5.7)	<0.001
MCH	29.76 (28.87, 30.53)	30.10 (29.13, 31.00)	0.3	29.67 (28.76, 30.58)	30.10 (29.15, 31.00)	0.11
MCHC	325 (10)	332 (12)	<0.001	325 (12)	332 (12)	0.015
MCV	91.9 (88.4, 93.4)	90.7 (88.0, 93.2)	0.3	91.2 (87.8, 93.8)	90.6 (87.8, 93.4)	0.9
PDW	11.73 (10.93, 12.70)	12.53 (11.32, 14.18)	0.017	11.02 (9.95, 12.90)	12.43 (11.35, 14.00)	0.005
RDW	13.40 (12.87, 14.21)	13.10 (12.65, 13.65)	0.076	13.60 (13.09, 14.42)	13.10 (12.70, 13.70)	0.008
ALT	14 (12, 25)	16 (12, 24)	0.6	14 (10, 17)	17 (12, 25)	0.064
AST	21 (18, 26)	21 (18, 26)	0.9	22 (17, 27)	21 (18, 26)	>0.9
UA	244 (219, 324)	297 (241, 364)	0.018	251 (206, 292)	296 (237, 364)	0.013
CHE	6795 (1846)	7463 (1801)	0.058	6522 (1830)	7492 (1784)	0.022
TP	66 (7)	68 (6)	0.072	67 (6)	68 (6)	0.4
Globulin	27.0 (24.2, 29.3)	27.5 (24.6, 30.8)	0.3	27.6 (25.2, 30.4)	27.6 (24.5, 30.7)	0.6
D-dimer	4.56 (2.46, 6.45)	0.85 (0.26, 3.15)	<0.001	5.34 (3.27, 7.57)	0.85 (0.25, 3.06)	<0.001
FIB	3.36 (2.82, 4.07)	2.75 (2.33, 3.32)	<0.001	3.08 (2.64, 3.75)	2.78 (2.30, 3.32)	0.031
APTT	29.7 (27.2, 34.8)	28.1 (25.1, 33.4)	0.2	28.3 (25.7, 33.8)	28.3 (25.1, 33.8)	0.8
TT	16.00 (14.98, 17.00)	16.70 (15.70, 18.10)	0.016	16.15 (15.48, 17.56)	16.70 (15.70, 18.10)	0.13
PT	11.60 (10.75, 13.09)	11.45 (10.60, 12.40)	0.092	11.77 (10.70, 12.89)	11.40 (10.60, 12.40)	0.4

[a] Median (IQR); Mean (SD).
[b] Wilcoxon rank sum test; Two Sample Student t-test.

8.2. Model development

Of the 31 covariates included, we identified 15 potential predictors in the univariate logistics regression analysis within the training cohort (Supplemental Materials, Table S2). In the multivariate analysis, a backward stepwise method was used to further select these predictive factors, resulting in 5 factors retained in the final model without multicollinearity issues. Specifically, the multivariate logistics regression analysis indicated that Age (OR:1.06; 95%CI: 1.03–1.09), MCHC (0.95; 95%CI: 0.92–0.98), D-dimer (1.18; 95%CI: 1.04–1.34), PDW (0.83; 95%CI: 0.68–0.99), and TT (0.85; 95%CI: 0.69–1.03) were the predictive factors for DVT after LEOS (Table 3).

8.3. Model performance and validation

A nomogram was constructed in accordance with the 5 predictive factors above (Fig. 2), with each predictor given a score based on the effect of each hospitalized patient; the total score was then calculated to obtain the risk probability. The C-index of the nomogram was 85.8 %. The AUC was refined to 85.9 % after 2000 stratified bootstrap replicates in the training cohort, and 85.7 % in the validation cohort (Fig. 3), indicating a satisfactory discriminative ability of the model. The ROC curve for each of the five predictors was provided in the supplemental material (Fig. S1), with Age and D-dimer obtaining the best predictive performance individually.

A good concordance between the observation and prediction was observed in the calibration curve in the training set, as well as in the validation set except there was a slight departure from the ideal line (Fig. 4). Accompanying with the Hosmer and Lemeshow test result, a p-value of 0.8791 suggested no evidence of poor fit.

Compared to two clinical schemas, intervention for all and intervention for none, the decision curve revealed that the DVT prediction nomogram would provide a positive net benefit when the threshold probability was between 0.01 and 0.16 (Fig. 5).

Table 3
Multivariate logistic regression analysis in patients undergoing lower extremity orthopedic surgery for DVT prediction.

Characteristic	OR[a]	95 % CI[1]	p-value
Age	1.06	1.03–1.09	<0.001
MCHC (g/L)	0.95	0.92–0.98	0.004
D-dimer (mg/L)	1.18	1.04–1.34	0.010
PDW (%)	0.83	0.68–0.99	0.055
TT (s)	0.85	0.69–1.03	0.12

[a] OR = Odds Ratio, CI = Confidence Interval.

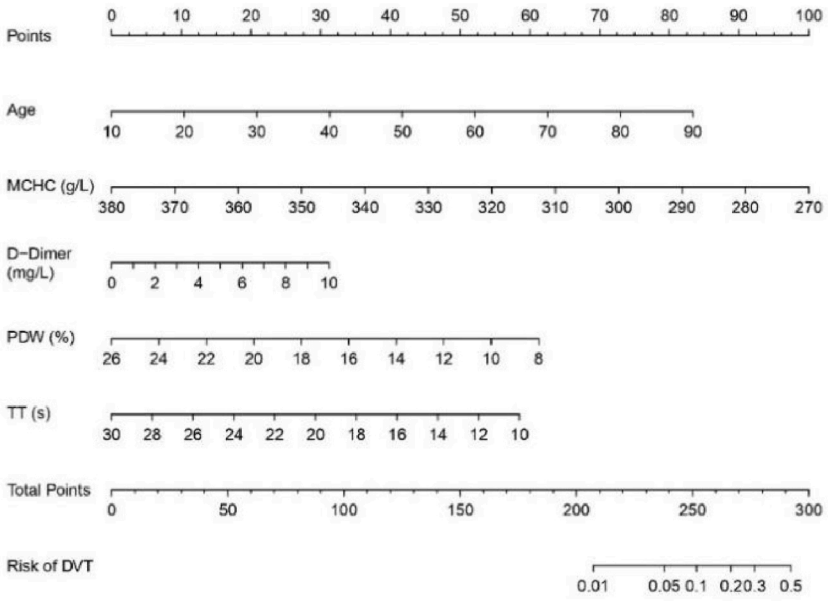


Fig. 2. The nomogram constructed in accordance with the five predictive factors. MCHC, mean corpuscular hemoglobin concentration; PDW, platelet distribution width; TT, thrombin time.

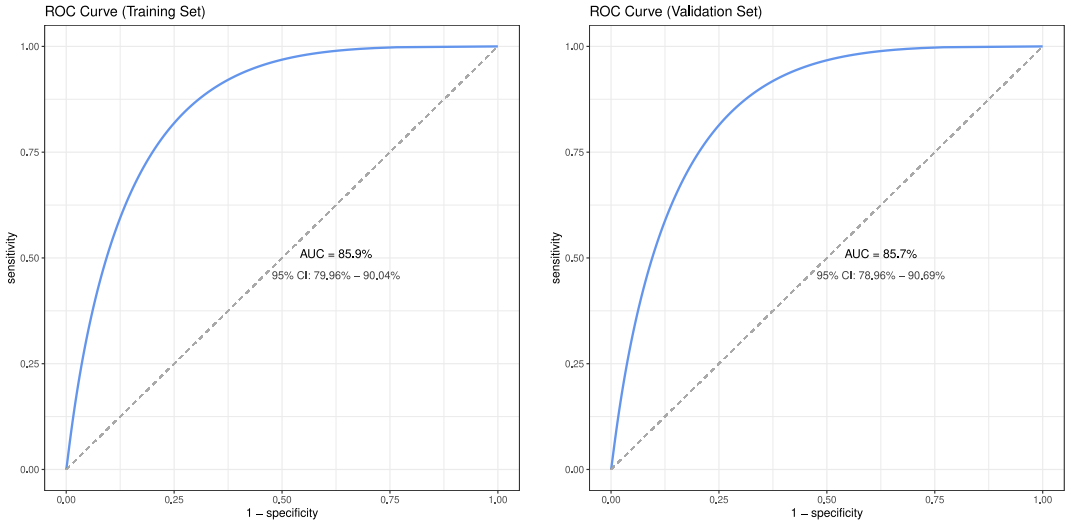


Fig. 3. The AUC in the training and validation cohort.

8.4. Sensitivity analysis

Using the dataset without imputation, the multivariate logistics regression analysis revealed similar findings and the results were fairly comparable (Supplemental Materials, Table S3). Therefore, the imputation values are robust, and the assumption of MAR is demonstrated to be valid.

9. Discussion

A total of 5773 patients with LEOS, approximately 1 % (52/5773) of whom had DVT, were selected in this study. After preliminary selected for 15 potential predictors of the 31 covariates and further selected by backward stepwise method in the training cohort of 4041, we finally retained 5 factors, including age, MCHC, D-dimer, PDW, and TT. The AUC was refined to 85.9 % in the training cohort, and 85.7 % in the validation cohort, showing a satisfactory discriminative ability of the model. Similar from the other reported models of DVT risk prediction, we considered it was because composite endpoint diagnosis of LEOS.

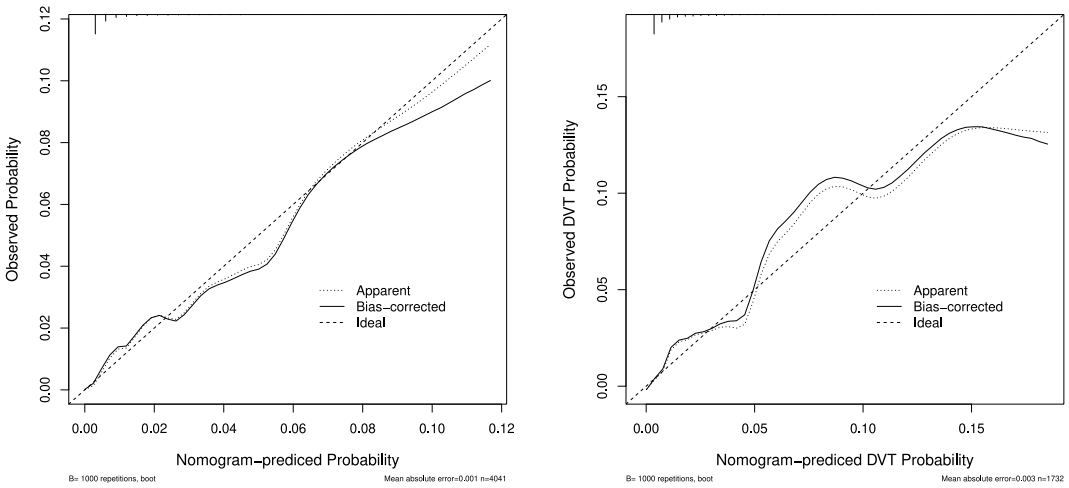


Fig. 4. The calibration curve in the training and validation set.

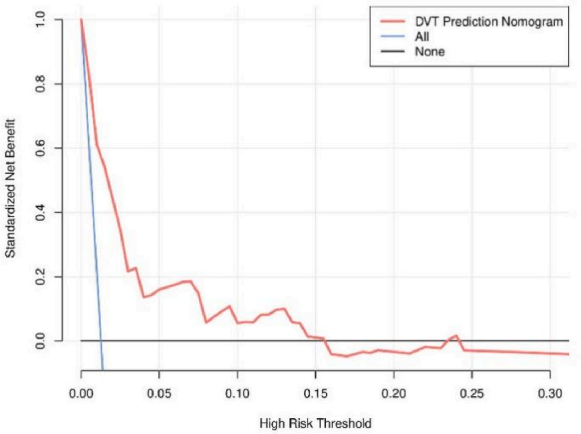


Fig. 5. The decision curve of DVT prediction nomogram.

In our study the median age of patients with or without DVT was 60 (IQR, 48–70) and 75 (67–82), mainly as the middle-old, which was corresponding with the discoveries that have been described in other studies [49–51]. One study indicated that in middle-aged population (40–44 years old), the incidence of DVT was 30 % per 1000 person-years(py), which was three times more frequent than PE; while the situation was the opposite in older patients of 80–84 years old [51]. However, people range from 70 to 74 years old was the smaller peak compared with that of 20–24 years old in Jiangxi province in China, probably because of the different exclusion criteria and geographic region [39]. Whether the young or the old, the predictor of the age influenced the incidences of DVT, and subsequently the construction of predictive models [36,37]. Because of the lower specificity of D-dimer testing in older population, age-dependent D-dimer cut-off values calculated by multiplying 10 times of the patients' age in patients with over 50 years old were applied in quite a lot of studies and usually performed better compared to traditional D-dimer of 500 μg/L [52–54]. In our study, the novel predictive model we have developed has exhibited a significant enhancement in accuracy, particularly when applied to distinct age cohorts, surpassing the performance metrics of its predecessors. This refinement has profound clinical implications, as it paves the way for the timely detection of individuals at elevated risk. Consequently, this advancement facilitates the implementation of targeted preventive interventions at a more advanced stage, potentially leading to a marked reduction in the incidence of adverse outcomes. The AUC of D-dimer alone was 0.776 (Fig. S1) with somewhat low specificity of 68.56 % (data not shown) in our study, partly because patients with DVT were the middle-old (median age of 60 years old). In our investigation, the comprehensive predictive model we have engineered has achieved AUC exceeding 0.85 in both the training and validation datasets. This performance is notably superior to the AUC values of less than 0.8 reported in other literature for differentiating between DVT and non-DVT patients [55,56]. These findings underscore the enhanced discriminatory capability of our model, suggesting a more refined approach to clinical diagnostics and patient risk stratification. Except for D-dimer, the quite strong predictor that has been described in enormous studies, other predictors in our study such as PDW and TT, were confirmed by different researches. For example, PDW was an independent predictor for postoperative DVT events of people experienced cervical carcinoma [57], and for preoperative DVT of ones followed foot who

experienced fractures [58]. As the same combination of D-Dimer with PDW, or TT in our study, it was considered to have a significant ability to distinguish between DVT and non-DVT in other results [59,60].

MCHC, as one of the common clinical indicators of coagulation function, was also illustrated as a predictor in our prediction model of DVT risk after LEOS. Interestingly, although several studies have confirmed the similar opinion [61,62], others supposed that MCHC was not significantly associated with DVT [63–65]. In the data from the MEGA study, a large population of 2473 patients with VTE and 2935 controls, showed MCHC was one of the risk factor of VTE [62]. Another study of malignancies with different thromboembolic conditions also illustrated that the level of MCHC affected physiological indicators in patients with and without DVT, respectively [61]. The opposite hypothesis indicated that even with the other decreased routine blood indicators, such as RBC count, and HCT levels, there was no significant differences with MCHC in the HIV negative-DVT [63,65], which was consistent with the findingin a prospective two-center cohort study evaluating the risk factor of lower extremity venous thrombosis after urologic surgeries [64]. This observation may underpin the physiological basis for their combined predictive utility in assessing the risk of DVT. Specifically, the interplay between D-dimer and platelets may be intrinsically linked to their synergistic roles in the coagulation cascade, a hypothesis that aligns with the clinical mechanisms observed in thrombogenesis. This insight offers a nuanced understanding of the molecular interactions that drive the pathogenesis of DVT, potentially informing more targeted and effective therapeutic strategies.

We have to say that the machine learning did offer practical value combined with traditional methods for screening DVT [66], and can improve the diagnostic process of suspected DVT patients, simultaneously reducing the healthcare resource expenditures [67]. The application of our predictive model holds the potential to enhance the clinical detection rate of DVT while simultaneously mitigating the occurrence of adverse events. This advancement is of significant importance for improving patient outcomes and reducing healthcare costs, aligning with the objectives of precision medicine and the pursuit of optimal clinical governance. Comparing effectiveness of machine learning methods for diagnosis of DVT, apart from their respective pitfalls compared to conventional methods, the best approaches were timely and effective [44,66,68,69]. Nevertheless, the poor data consistency and completeness, as well as difficulties in data mining in medical data of Chinese hospitals limited the application of AI algorithms to prediction model constructions [70,71]. By targeting this problem, we believed that this step of data cleaning needs to be optimized first, in which MICE made the critical role. Methodologically, multivariate imputation (MI) was recommended superiorly to handle such missing data in clinical prediction models [72,73]. Kristel et al. suggested MI was preferred when applying clinical prediction models. The authors developed and validated a model to predict the risk of DVT in 1295 and 532 primary care patients, respectively, and found MI led to calibration intercepts closest to the true value in all three situations of missing an important predictor (D-dimer test), or a weaker predictor (difference in calf circumference), or both two [74]. Furthermore, as a principled method of dealing with missing data [75], MICE was proved to perform quite well in DVT-related studies [76,77].

9.1. Strengths and limitations

Although many studies related to the modeling of DVT after THR or TKR surgery have been reported, this study is the first to address the prediction of DVT risk in patients after LEOS. Given the clinical data was only collected from one single center, the model could be better validated and generalized if there is a multicenter study in the future. In addition, this study focused only on the postoperative period, further study incorporated with preoperative DVT risk prediction to make it a better estimation during the perioperative period would be more meaningful for the management of LEOS.

10. Conclusion

In this study, we constructed a prediction model for DVT risk after LEOS based on MICE consisting of 5 key predictors, age, MCHC, D-dimer, PDW, and TT. It fills the gap of prediction DVT risk in the postoperative period of LEOS, our refinements have not only augmented the accuracy of diagnostic procedures but also harbor the potential to exert a salutary impact on health economics by diminishing the incidence of unwarranted D-dimer testing and concomitant medical expenditures. Subsequent research endeavors should endeavor to probe the model's applicability and cost-effectiveness across a spectrum of populations, thereby elucidating its broader implications for clinical practice and healthcare policy.

Data sharing statement

All data collected for the study are presented in the manuscript and its supplements.

Data sharing statement

All data collected for the study are presented in the manuscript and its supplements.

CRediT authorship contribution statement

Jiannan Zhang: Conceptualization. **Yang Shao:** Data curation. **Hongmei Zhou:** Investigation. **Ronghua Li:** Data curation. **Jie Xu:** Investigation, Conceptualization. **Zhongzhou Xiao:** Writing – original draft, Investigation, Formal analysis. **Lu Lu:** Writing – review & editing, Writing – original draft, Investigation. **Liangyu Cai:** Writing – review & editing, Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This study was supported by the fund of Wuxi Health Bureau (No. MS201935).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29517.

References

[1] R.J. Weiss, et al., Orthopaedic surgery of the lower limbs in 49 802 rheumatoid arthritis patients: results from the Swedish National Inpatient Registry during 1987 to 2001, Ann. Rheum. Dis. 65 (3) (2006) 335–341.
[2] H. Doll, et al., Evaluation of the measurement properties of four performance outcome measures in patients with elective hip replacements, elective knee replacements, or hip fractures, Value Health 21 (9) (2018) 1104–1114.
[3] F. Piovella, et al., Deep-vein thrombosis rates after major orthopedic surgery in Asia. An epidemiological study based on postoperative screening with centrally adjudicated bilateral venography, J. Thromb. Haemostasis 3 (12) (2005) 2664–2670.
[4] C. Wang, et al., Simplified Chinese version of hip and knee replacement expectations surveys in patients with osteoarthritis and ankylosing spondylitis: cross-cultural adaptation, validation and reliability, BMC Muscoskel. Disord. 19 (1) (2018) 1–10.
[5] F.S. Drescher, et al., Aspirin versus anticoagulation for prevention of venous thromboembolism major lower extremity orthopedic surgery: a systematic review and meta-analysis, J. Hosp. Med. 9 (9) (2014) 579–585.
[6] O.o.t.S. General, L. National Heart, B. Institute, The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism, 2008.
[7] W.H. Geerts, et al., Prevention of venous thromboembolism: the seventh ACCP conference on antithrombotic and thrombolytic therapy, Chest 126 (3) (2004) 338S–400S.
[8] A. Leizorovicz, et al., Epidemiology of post-operative venous thromboembolismin asian countries, Int. J. Angiol. 13 (3) (2004) 101–108.
[9] S.Y. Lee, et al., Incidence of deep vein thrombosis after major lower limb orthopedic surgery: analysis of a nationwide claim registry, Yonsei Med. J. 56 (1) (2015) 139–145.
[10] H. Bo, et al., Assessing the risk for development of deep vein thrombosis among Chinese patients using the 2010 Caprini risk assessment model: a prospective multicenter study, J. Atherosclerosis Thromb. 27 (8) (2020) 801–808.
[11] T.L. Ortel, et al., American Society of Hematology 2020 guidelines for management of venous thromboembolism: treatment of deep vein thrombosis and pulmonary embolism, Blood advances 4 (19) (2020) 4693–4738.
[12] N.G.C. Uk, Venous Thromboembolism in over 16s: Reducing the Risk of Hospital-Acquired Deep Vein Thrombosis or Pulmonary Embolism, 2018.
[13] A. Tanavalee, Should Prophylaxis of Venous Thromboembolism in Asian Patients Undergoing Knee and Hip Arthroplasty and Hip Fracture Surgery Be an Issue? Springer, 2021, pp. 1–2.
[14] A. Nicolaides, et al., Prevention and treatment of venous thromboembolism: international consensus statement (guidelines according to scientific evidence), Clin. Appl. Thromb. Hemost. 19 (2) (2013) 116–118.
[15] S. Ngarmukos, et al., Asia-Pacific venous thromboembolism consensus in knee and hip arthroplasty and hip fracture surgery: Part 1. Diagnosis and risk factors, Knee Surgery & Related Research 33 (1) (2021) 1–13.
[16] J.B. Brill, et al., The rate of deep vein thrombosis doubles in trauma patients with hypercoagulable thromboelastography, J. Trauma Acute Care Surg. 83 (3) (2017) 413–419.
[17] S. Decker, M. Weaver, Deep venous thrombosis following different isolated lower extremity fractures: what is known about prevalences, locations, risk factors and prophylaxis? Eur. J. Trauma Emerg. Surg. 39 (6) (2013) 591–598.
[18] W.F. Hefley Jr, C.L. Nelson, C.L. Puskarich-May, Effect of delayed admission to the hospital on the preoperative prevalence of deep-vein thrombosis associated with fractures about the hip, JBJS 78 (4) (1996) 581–583.
[19] H.R. Zahn, J.A. Skinner, M. Porteous, The preoperative prevalence of deep vein thrombosis in patients with femoral neck fractures and delayed operation, Injury 30 (9) (1999) 605–607.
[20] S.-W. Qu, et al., Deep vein thrombosis in the uninjured lower extremity: a retrospective study of 1454 patients with lower extremity fractures, Clin. Appl. Thromb. Hemost. 27 (2021) 1076029620986862.
[21] C. Mok, et al., The incidence of deep vein thrombosis in Hong Kong Chinese after hip surgery for fracture of the proximal femur, Br. J. Surg. 66 (9) (1979) 640–642.
[22] K. Song, et al., The preoperative incidence of deep vein thrombosis (DVT) and its correlation with postoperative DVT in patients undergoing elective surgery for femoral neck fractures, Arch. Orthop. Trauma Surg. 136 (10) (2016) 1459–1464.
[23] G. Piazza, S.Z. Goldhaber, Chronic thromboembolic pulmonary hypertension, N. Engl. J. Med. 364 (4) (2011) 351–360.
[24] S.R. Kahn, The post-thrombotic syndrome, Hematology 2010, the American Society of Hematology Education Program Book 2010 (1) (2010) 216–220.
[25] S.-H. Liu, et al., An examination system to detect deep vein thrombosis of a lower limb using light reflection rheography, Sensors 21 (7) (2021) 2446.
[26] D. Scarvelis, P.S. Wells, Diagnosis and treatment of deep-vein thrombosis, CMAJ (Can. Med. Assoc. J.) 175 (9) (2006) 1087–1092.
[27] G.H. Guyatt, et al., Methodology for the development of antithrombotic therapy and prevention of thrombosis guidelines: antithrombotic therapy and prevention of thrombosis: American college of chest physicians evidence-based clinical practice guidelines, Chest 141 (2) (2012) 53S–70S.
[28] N. Perone, H. Bounameaux, A. Perrier, Comparison of four strategies for diagnosing deep vein thrombosis: a cost-effectiveness analysis, Am. J. Med. 110 (1) (2001) 33–40.
[29] C.S. Landefeld, Noninvasive diagnosis of deep vein thrombosis, JAMA 300 (14) (2008) 1696–1697.
[30] D.G. Fraser, et al., Diagnosis of lower-limb deep venous thrombosis: a prospective blinded study of magnetic resonance direct thrombus imaging, Ann. Intern. Med. 136 (2) (2002) 89–98.
[31] L.-Y. Chong, et al., Management of venous thromboembolic diseases and the role of thrombophilia testing: summary of NICE guidance, Bmj (2012) 344.
[32] P.S. Wells, et al., Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis, N. Engl. J. Med. 349 (13) (2003) 1227–1235.
[33] P.S. Wells, et al., Value of assessment of pretest probability of deep-vein thrombosis in clinical management, Lancet 350 (9094) (1997) 1795–1798.
[34] S. Jin, et al., Machine learning predicts cancer-associated deep vein thrombosis using clinically available variables, Int. J. Med. Inf. 161 (2022) 104733.

[35] S.M. Bates, et al., Diagnosis of DVT: antithrombotic therapy and prevention of thrombosis: American College of Chest Physicians evidence-based clinical practice guidelines, Chest 141 (2) (2012) e351S–e418S.
[36] H. Liu, et al., Prediction of venous thromboembolism with machine learning techniques in young-middle-aged inpatients, Sci. Rep. 11 (1) (2021) 1–12.
[37] B. Linnemann, et al., Prevalence of established risk factors for venous thromboembolism according to age, J. Vasc. Surg.: Venous and Lymphatic Disorders 2 (2) (2014) 131–139.
[38] M. Ljungqvist, et al., Risk factors for recurrent venous thromboembolism in young and middle-aged women, Thromb. Res. 133 (5) (2014) 762–767.
[39] F. Chen, J.X. Xiong, W.M. Zhou, Differences in limb, age and sex of Chinese deep vein thrombosis patients, Phlebology 30 (4) (2015) 242–248.
[40] S.R. Kahn, S. Shivakumar, What's new in VTE risk and prevention in orthopedic surgery, Research and practice in thrombosis and haemostasis 4 (3) (2020) 366–376.
[41] E. Rasouli Dezfouli, et al., A machine learning framework for assessing the risk of venous thromboembolism in patients undergoing hip or knee replacement, Journal of Healthcare Informatics Research (2022) 1–19.
[42] S.K. Devana, et al., A novel, potentially universal machine learning algorithm to predict complications in total knee arthroplasty, Arthroplasty today 10 (2021) 135–143.
[43] K.N. Kunze, et al., Development and internal validation of machine learning algorithms for predicting complications after primary total hip arthroplasty, Arch. Orthop. Trauma Surg. (2022) 1–8.
[44] Y. Chen, Y. Jiang, Construction of prediction model of deep vein thrombosis risk after total knee arthroplasty based on XGBoost algorithm, Comput. Math. Methods Med. (2022) 2022.
[45] Z. Lin, et al., Nomogram for predicting deep venous thrombosis in lower extremity fractures, BioMed Res. Int. 2021 (2021).
[46] J. Ma, et al., Incidence and risk factors predicting deep venous thrombosis of lower extremity following spinal fractures, Sci. Rep. 11 (1) (2021) 1–8.
[47] R.D. Riley, et al., Calculating the sample size required for developing a clinical prediction model, Bmj 368 (2020).
[48] T.P.C.o.p.v.d.o.t.C.A.o.I. Medicine, Diagnostic and efficacy criteria for lower extremity deep vein thrombosis (revised 2015), Journal of the Chinese academy of integrative medicine 22 (5) (2016) 2.
[49] M. Engbers, A. van Hylckama Vlieg, F. Rosendaal, Venous thrombosis in the elderly: incidence, risk factors and risk groups, J. Thromb. Haemostasis 8 (10) (2010) 2105–2112.
[50] W. Ageno, et al., Risk factors for venous thromboembolism in the elderly: results of the master registry, Blood Coagul. Fibrinolysis 19 (7) (2008) 663–667.
[51] M.D. Silverstein, et al., Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study, Arch. Intern. Med. 158 (6) (1998) 585–593.
[52] J.X. Wu, et al., Performance of age-adjusted D-dimer values for predicting DVT before the knee and hip arthroplasty, J. Orthop. Surg. Res. 16 (1) (2021) 1–6.
[53] K. Broen, B. Scholtes, R. Vossen, Predicting the need for further thrombosis diagnostics in suspected DVT is increased by using age adjusted D-dimer values, Thromb. Res. 145 (2016) 107–108.
[54] R.A. Douma, et al., Age-dependent D-dimer cut-off value increases the number of older patients in whom deep vein thrombosis can be safely excluded, Haematologica 97 (10) (2012) 1507–1513.
[55] J.X. Wu, et al., Performance of age-adjusted D-dimer values for predicting DVT before the knee and hip arthroplasty, J. Orthop. Surg. Res. 16 (2021) 1–6.
[56] C. Wei, et al., Comparison of different machine learning classification models for predicting deep vein thrombosis in lower extremity fractures, Sci. Rep. 14 (1) (2024) 6901.
[57] N. Li, et al., Combination of preoperative D-dimer and platelet distribution width predicts postoperative deep venous thrombosis in patients with cervical carcinoma, Asian Pac. J. Cancer Prev. APJCP 20 (4) (2019) 1025.
[58] J. Ma, et al., Incidence and hematological biomarkers associated with preoperative deep venous thrombosis following foot fractures, Foot Ankle Int. 41 (12) (2020) 1563–1570.
[59] Y. Zhang, et al., The value of platelet parameters, fibrinogen and D-dimer in the diagnosis of deep vein thrombosis, Med. J. Chin. Peoples Lib. Army 46 (9) (2021) 906–911.
[60] X. Wang, et al., Prevalence of preoperative Deep Venous Thrombosis (DVT) following elderly intertrochanteric fractures and development of a risk prediction model, BMC Muscoskel. Disord. 23 (1) (2022) 1–9.
[61] Y. Li, Clinical Features of Malignant Tumors Associated with Different Thromboembolic Disorders, Henan University, 2020.
[62] S.M. Rezende, et al., Hematologic variables and venous thrombosis: red cell distribution width and blood monocyte count are associated with an increased risk, Haematologica 99 (1) (2014) 194.
[63] B.S. Jackson, J. Nunes Goncalves, E. Pretorius, Comparison of pathological clotting using haematological, functional and morphological investigations in HIV-positive and HIV-negative patients with deep vein thrombosis, Retrovirology 17 (1) (2020) 1–13.
[64] G. Tang, et al., Evaluation and analysis of incidence and risk factors of lower extremity venous thrombosis after urologic surgeries: a prospective two-center cohort study using LASSO-logistic regression, Int. J. Surg. 89 (2021) 105948.
[65] B.S. Jackson, J. Nunes, E. Pretorius, Haematological Parameters and Pathological Clotting in Deep Vein Thrombosis and Patients with HIV, 2019.
[66] R. Sorano, L.V. Magnusson, K. Abbas, Comparing effectiveness of machine learning methods for diagnosis of deep vein thrombosis, in: International Conference on Computational Science and its Applications, Springer, 2022.
[67] L. Luo, et al., Cost-effective machine learning based clinical pre-test probability strategy for DVT diagnosis in neurological intensive care unit, Clin. Appl. Thromb. Hemost. 27 (2021) 10760296211008650.
[68] E.E. Contreras-Luján, et al., Evaluation of machine learning algorithms for early diagnosis of deep venous thrombosis, Math. Comput. Appl. 27 (2) (2022) 24.
[69] G. Peng, et al., Development and prospective validation of a novel risk score for predicting the risk of lower extremity deep vein thrombosis among multiple trauma patients, Thromb. Res. 201 (2021) 116–122.
[70] S. Tian, et al., Smart healthcare: making medical care more intelligent, Global Health Journal 3 (3) (2019) 62–65.
[71] W. Sun, et al., Data processing and text mining technologies on electronic medical records: a review, Journal of healthcare engineering 2018 (2018).
[72] K.G. Moons, et al., Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration, Ann. Intern. Med. 162 (1) (2015) W1–W73.
[73] A. Rabinovich, et al., Development of a clinical prediction model for the postthrombotic syndrome in a prospective cohort of patients with proximal deep vein thrombosis, J. Thromb. Haemostasis 16 (2) (2018) 262–270.
[74] K.J. Janssen, et al., Dealing with missing predictor values when applying clinical prediction models, Clin. Chem. 55 (5) (2009) 994–1001.
[75] M.J. Azur, et al., Multiple imputation by chained equations: what is it and how does it work? Int. J. Methods Psychiatr. Res. 20 (1) (2011) 40–49.
[76] G. Cafri, et al., Comparative effectiveness and safety of drug prophylaxis for prevention of venous thromboembolism after total knee arthroplasty, J. Arthroplasty 32 (11) (2017) 3524–3528. e1.
[77] G.J. Van der Heijden, et al., Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example, J. Clin. Epidemiol. 59 (10) (2006) 1102–1109.