

Anti-factor Xa during unfractionated heparin therapy in critically ill patients: Development of prediction models using machine learning

Boris Delange^{1,*} , Guillaume Bouzillé², Pauline Guillot¹, Anaëlle Bichon¹, Océane Bernard de Lajartre¹, Isabelle Gouin³, Yoann Launey⁴, Alexandre Mansour^{3,4}, Mathieu Lesouhaitier¹, Jean-Marc Tadié^{1,5}, Arnaud Gacouin^{1,5}, Marc Cuggia² and Adel Maamar¹

Abstract

Objective: Unfractionated heparin (UFH) is a widely used therapy in intensive care units (ICUs) and is associated with an increased risk of serious adverse events or death if the therapeutic target is not reached quickly. Adjusting UFH dosage is challenging, and no reliable algorithms exist for predicting anti-Xa levels in ICUs. This study aimed to develop and evaluate machine learning algorithms to predict anti-Xa levels during UFH therapy, helping clinicians optimize dosing.

Methods: This single-center retrospective cohort study was conducted using Rennes University Hospital's clinical data warehouse from December 21, 2019 to November 22, 2021. Critically ill patients ≥ 18 years on UFH, without other anticoagulants and complete data, were included. Anti-Xa levels were classified as infra-therapeutic (<0.3), therapeutic (0.3 – 0.7), or supra-therapeutic (>0.7). Models incorporated UFH rate, bolus, prior anti-Xa, kidney function, inflammation, volemic state, extracorporeal membrane oxygenation, and bilirubinemia. Performance was assessed using the area under the receiver operating characteristic curve (AUROC), area under the precision-recall curve (AUPRC), sensitivity, and specificity.

Results: A total of 3790 anti-Xa intervals, corresponding to 211 patients, were included in the study. Out of several machine learning algorithms, random forest achieved the best results with an AUROC score of 0.80 [0.77;0.83], an AUPRC score of 0.61 [0.58;0.65], sensitivity of 0.56 [0.53;0.59] and specificity of 0.82 [0.82;0.82].

Conclusion: In this cohort study, machine learning-based prediction models achieved good performance for predicting anti-Xa results during UFH therapy in an ICU setting. Further validation with prospective multicenter data is needed in order to confirm the model's generalizability and support its integration into clinical practice to assist clinicians in selecting the optimal heparin dose.

Keywords

Unfractionated heparin, intensive care unit, machine learning, prediction models, anti-Xa

Submission date: 11 July 2024; Acceptance date: 19 November 2024

¹Service de Maladies Infectieuses et Réanimation Médicale, Hôpital Pontchaillou, Université de Rennes, Rennes, France

²CHU Rennes, INSERM, LTSI-UMR 1099, Université de Rennes, Rennes, France

³Laboratoire d'Hématologie-Hémostase, Centre Hospitalo-Universitaire de Rennes, Rennes, France

⁴Département Anesthésie Réanimation et Médecine Périopératoire, Centre Hospitalier Universitaire Rennes, Rennes, France

⁵Faculté de Médecine, Université de Rennes 1, Unité INSERM CIC 1414, IFR 140, Rennes, France

*Current affiliation: INSERM, LTSI-UMR 1099, Université de Rennes, Rennes, France.

Corresponding author:

Boris Delange, CHU Rennes, INSERM, LTSI-UMR 1099, Université de Rennes, 35000 Rennes, France.

Email: boris.delange@univ-rennes.fr



Background

Continuous unfractionated heparin (UFH) is a widely used therapy for thromboembolic events, anticoagulation of extracorporeal circuits, atrial fibrillation, arterial thrombosis and bridging for long-term oral anticoagulation. Despite the increasing use of low-molecular-weight heparin and direct oral anticoagulants, UFH is still widely used in critically ill patients because of its shorter half-life, enabling the therapy to be stopped quickly in the case of a hemorrhagic event.¹ Thus, reaching the therapeutic range is necessary to achieve the therapeutic effect as quickly as possible and limit the onset or progression of thrombosis, while also limiting the risk of bleeding from overdose. In fact, the risk of bleeding is higher for supra-therapeutic ranges of aPTT (activated partial thromboplastin time),² while the mortality rate of pulmonary embolism and the recurrence of venous thromboembolic events decrease when the therapeutic target of aPTT is reached quickly.^{3,4} Consequently, treatment with UFH relies on frequent therapeutic monitoring and subsequent dose adjustments to ensure optimal efficacy and patient safety. In clinical practice, UFH therapy is typically started with a bolus dose followed by continuous infusion.⁵ The continuous infusion rate is adjusted for each biological monitoring result to reach the therapeutic range.

UFH acts by binding to antithrombin (AT), enhancing its inhibitory activity primarily against factors IIa and Xa. As previously described,^{6–12} the pharmacokinetic (PK) properties of UFH are influenced by kidney function, liver function, inflammation, and plasma volume.

To assess UFH's antithrombotic effect, intensivists widely use an anti-factor Xa assay (anti-Xa), which is less influenced by biological factors, inflammation and infectious conditions than aPTT. It is also associated with less variability in UFH dosage and so requires fewer monitoring tests,¹³ and demonstrates faster achievement of therapeutic range.¹⁴

To date, nomograms have been applied, taking the weights of several factors into consideration to reach the aPTT target.⁵ However, these nomograms have shown poor accuracy in critically ill patients. To address this issue, machine learning algorithms were developed to include more factors, with promising results,^{15–17} but focused solely on aPTT results.

Therefore, a machine learning approach might be helpful to predict anti-Xa results more accurately and help clinicians to adjust UFH dosage.

In this study, we sought to develop an explainable machine learning prediction model for anti-Xa assay results, after initiation of UFH, consisting of all parameters known to influence results and readily available in intensive care unit (ICU) daily practice. We verified the model's prediction performance by applying these models to test data.

Methods

Study population

Data was extracted retrospectively from the clinical data warehouse¹⁸ of a tertiary teaching hospital. Patients were enrolled if they stayed in medical, surgical or cardiac surgery ICUs from December 21, 2019 to November 22, 2021. Anti-Xa results from patients under 18 years ($n=60$), who received protamine sulfate in the 72 hours prior to the anti-Xa assay, subcutaneous unfractionated heparin or low-molecular-weight heparin in the previous 48 hours, direct oral anticoagulants or antithrombin in the previous 7 days ($n=1061$), with an interval between two anti-Xa assays of less than 6 h or more than 48 h ($n=760$), with missing data on anti-Xa results ($n=670$) or for at least one of the features after imputation ($n=1147$) were excluded, leaving 3790 anti-Xa results included in the analysis. The study was approved by the Rennes University Hospital institutional review board (No. 23.03) and followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (**TRIPOD**) reporting guidelines.¹⁹

Outcome

The predicted outcome was the anti-Xa assay result, considered sub-therapeutic if less than 0.3, normal between 0.3 and 0.7 or supra-therapeutic if greater than 0.7.

Features for prediction models

Anti-Xa activity was measured on fresh plasma, collected in 0.109 M citrate tubes, using a commercially available chromogenic anti-FXa assay (STA Liquid Anti-Xa[®] from Stago, with no dextran) and dedicated calibrators and controls, on STA R Max 2[®] (Stago).

We extracted data at intervals between two anti-Xa assays spaced 6 to 48 h apart. We selected 16 features as explanatory variables, which are routinely collected for ICU patients: demographic features with age, gender and body mass index (BMI) on admission, use of extracorporeal membrane oxygenation (ECMO), evaluation of kidney function with urine output, creatinemia and use of renal replacement therapy (RRT) during the interval, evaluation of inflammation with fibrinogenemia, evaluation of distribution volume with hemoglobinemia and hematocrit, bilirubinemia, unfractionated heparin dose administered during the previous and current intervals, unfractionated heparin bolus administered during the interval, anti-Xa result at the interval start and interval duration.

Data preprocessing

Outlier data boundaries were evaluated for each feature by two clinicians (BD, AM) (Supplemental Table 1). Urine output and UFH boli were summed over the intervals, the other continuous variables were meaned over the intervals. Patients with missing data for demographic features and unfractionated heparin rates were excluded for the intervals concerned. Missing data on laboratory tests results were imputed using linear interpolation, which was chosen for its ability to maintain temporal continuity and handle time-series data where trends are typically gradual. More complex methods, such as k-nearest neighbors, were not used due to the relatively small dataset size and the risk of introducing biased data. Continuous variables were normalized, categorical variables were transformed with one-hot encoding, and right-skewed distributed variables were log-transformed and normalized. Examples of data transformations are shown in Supplemental Figure 1a and b.

Machine learning algorithms

The following six machine learning algorithms were used to build the classification models: random forest, linear support vector machine (SVM), radial basis function kernel support vector machine (RBF SVM), eXtreme gradient boosting trees (XGB), neural network and penalized logistic regression.

A decision tree is an algorithm that splits data into multiple subsets, creating a tree-like structure whose nodes split the data using feature cutoffs and leaves contain predictions. Random forest uses a system of voting between an ensemble of decision trees, where each decision tree is trained on different subsets with different features. Like random forest, gradient boosting trees consist of an ensemble of decision trees, but while random forest trains each tree independently, gradient boosting trains the trees sequentially, minimizing the errors of the previous ones. SVM algorithms split the data with a hyperplane with a maximum margin between the data points and the hyperplane. The best hyperplane is the decision boundary. For non-linear problems, certain transformations, called kernels, are used to add dimensions and solve them. Radial basis function is one possible kernel used for SVM. Penalized logistic regression is an algorithm that assigns coefficients to each feature and makes predictions through a sigmoid function. Neural networks, modeled loosely on the human brain, create a network with multiple nodes that are densely interconnected and arranged into layers. Each node contains a weight value and is activated or disabled using an activation function. Lastly, in the final layer the “most activated” node is the model’s prediction.

The hyperparameters used in the grid search are shown in Supplemental Table 2. Latin hypercube sampling, a space-filling method, was used for the grid search. This method was chosen for its ability to efficiently explore a high-dimensional hyperparameter space, compared to traditional grid search, which can be computationally expensive and redundant in sampling similar points. An example is shown in Supplemental Figure 2. The default values were used for the other parameters. Ten-fold stratified cross-validation was performed to assess model performance and select the best hyperparameters.

Training, cross-validation and testing

We used 70% of the data to train the classification models through 10-fold stratified cross-validation. The remaining 30% of data was reserved to test the predictive ability of the resulting classification models. The hyperparameters were selected based on the best area under the receiver operating characteristic curve (AUROC) score. The trained classifier models calculated three class probabilities for each patient (sub-therapeutic, normal, and supra-therapeutic classes). Our data set was imbalanced owing to the relative rarity of the supra-therapeutic class. To address this issue, case weights were assigned to the observations²⁰ to give equal total weight to each class. Case weights were available only for penalized logistic regression, random forest and XGB algorithms. We did not use any method to calibrate our models.

Performance evaluation

To evaluate the models’ discrimination performance, we calculated the AUROC, the area under the precision-recall curve (AUPRC), accuracy, F1 score, sensitivity, and specificity. Given the multiclass issue, overall AUROC values were calculated using the Hand and Till method,²¹ and overall AUPRC values were calculated using a macro averaging method. Calibration plots were used to evaluate the models’ calibration performance.

Statistical analysis

Continuous variables are presented as medians [interquartile ranges]. Categorical variables are presented as numbers [percentages]. Wilcoxon tests were used to compare numerical variables, and Fisher’s exact tests were used to compare categorical variables. The statistical analyses were performed using R version 3.6.0 (The R Group for Statistical Computing, Vienna, Austria). All *p* values of less than 0.05 were considered significant.

Table 1. Baseline characteristics of the study cohort.

Feature	Whole cohort (n = 615)	Included stays (n = 286)
Gender		
Male	417 (68)	224 (78)
Female	198 (32)	62 (22)
Age-years	63.0 (51.0–71.0)	65.0 (52.0–71.8)
BMI-kg.m ⁻²	26.5 (23.1–30.3)	26.9 (23.7–30.9)
Unit		
Surgical ICU	304 (49)	114 (40)
Medical ICU	190 (31)	104 (36)
Cardiac surgery ICU	121 (20)	68 (24)
Chronic kidney disease	121 (20)	38 (13)
Ischaemic heart disease	108 (18)	58 (20)
Diabetes mellitus	105 (17)	62 (22)
Atrial fibrillation	86 (14)	46 (16)
Chronic dialysis	72 (12)	16 (5.6)
COPD	60 (9.8)	25 (8.7)
Cirrhosis	30 (4.9)	21 (7.3)
SOFA-respiration (/4)	1 (0–2)	2 (0–3)
SOFA-coagulation (/4)	0 (0–1)	0 (0–1)
SOFA-liver (/4)	0 (0–0)	0 (0–0)
SOFA-circulatory (/4)	4 (0–4)	4 (1–4)
SOFA-central nervous system (/4)	4 (1–4)	4 (4–4)
SOFA-kidney (/4)	1 (0–3)	1 (0–2)
SOFA-total (/24)	11 (8–13)	11 (8–13)
Mechanical ventilation at ICU admission	333 (54)	184 (64)
Mechanical ventilation during ICU stay	478 (78)	263 (92)

(continued)

Table 1. Continued.

Feature	Whole cohort (n = 615)	Included stays (n = 286)
ECMO at admission	52 (8.5)	41 (14)
ECMO during ICU stay	89 (14)	69 (24)
Length of stay-days	10.1 (4.1–21.4)	16.3 (8.6–31.1)
Number of UFH episodes	1 (1–1)	1.0 (1–1)
Number of anti-Xa assays	6 (2–14)	7 (3–16)

Data are median (Q1–Q3) for continuous variables and number (%) for categorical variables.

Definition of abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit; SOFA, sequential organ failure assessment; ECMO, extracorporeal membrane oxygenation; UFH, unfractionated heparin

Results

Baseline characteristics

The study cohort included 3790 anti-Xa intervals corresponding to 211 patients (165 male patients [78%]; median [Q1–Q3] age 66.0 [54.0–72.0] years; median BMI 27.6 [23.8–31.2] kg.m⁻²) and 286 ICU stays (Table 1, Supplemental Table 3 and Figure 1). During their ICU stay, the median admission sequential organ failure assessment (SOFA) score was 11 [8–13], 92% patients were on mechanical ventilation and 24% patients were treated with ECMO. The length of ICU stay was 16.3 [8.6–31.1] days. All the anti-Xa intervals were divided into a training set (2652 intervals, corresponding to 279 ICU stays) and a test set (1138 intervals, corresponding to 224 ICU stays) by random allocation. Table 2 and Supplemental Table 4 show the distribution of features between the two sets. No statistically significant difference was found.

Comparison of machine learning algorithms in cross-validation

Figures 2 and 3 show the receiver operating characteristic (ROC) and precision-recall (PR) curves for the six machine learning algorithms. Out of all the algorithms, XGB and penalized logistic regression achieved the highest mean AUC score in the ROC curve analysis (XGB 0.79 [0.76–0.82]; penalized logistic regression 0.79 [0.77–0.81]), and penalized logistic regression, random forest and neural network achieved the highest mean AUC value in the PR curve (penalized logistic regression 0.57 [0.54–0.60]; random forest 0.57 [0.55–0.59]; neural network 0.57 [0.55–0.58]).

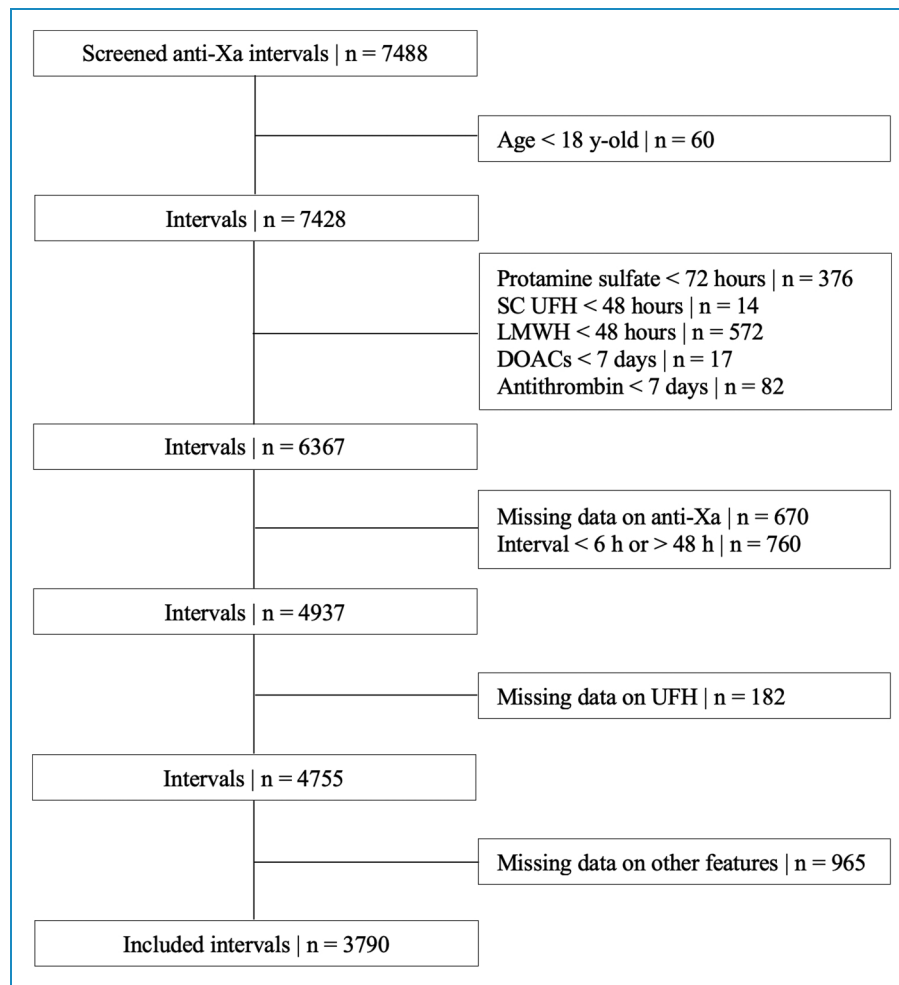


Figure 1. Diagram flowchart. SC UFH, subcutaneous unfractionated heparin; LMWH, low-molecular weight heparin; DOACs, direct oral anticoagulants; UFH, unfractionated heparin.

Comparison of machine learning algorithms in the test set

Out of all the algorithms, random forest and XGB achieved the highest mean AUC in the ROC curve (random forest 0.80 [0.77–0.83]; XGB 0.80 [0.76–0.83]), and random forest achieved the best AUC in the PR curve (0.61 [0.58–0.65]). Neural network achieved the best accuracy (0.74 [0.71–0.76]) and the best F1 score (0.76 [0.73–0.78]). Supplemental Table 5 shows the contingency tables, Supplemental Table 6 shows the global metrics, Supplemental Tables 7 and 8 show the AUROC and AUPRC of the models.

Evaluation of calibration

Supplemental Figure 3a and b shows the calibration plots for the six machine learning algorithms. Overall, calibration is good in the *Infra-therapeutic vs. others* and *Normal vs. others* outcomes, while low for the *Supra-therapeutic vs. others* outcome.

Feature importance

Feature importance was available only for the penalized logistic regression, random forest and XGB algorithms. The most important features are the UFH rate during the current interval, the UFH rate before the current interval and the start-of-interval anti-Xa (Supplemental Figure 4). The random forest model uses the different features in a balanced way.

Distribution of anti-Xa results

Supplemental Figure 5 shows the distribution of anti-Xa results in the whole cohort when the anti-Xa target is between 0.3 and 0.7. Forty-nine percent of anti-Xa results are under 0.3, classifying them as *infra-therapeutic*. Supplemental Figure 6 shows the distribution of the first anti-Xa result in the whole cohort when the anti-Xa target is between 0.3 and 0.7. Fifty-four percent of anti-Xa results are under 0.3.

Table 2. Comparison of features between train and test sets.

Feature	Train set (n = 279)	Test set (n = 224)	p-Value
Gender			0.91
Male	219 (78)	176 (79)	
Female	60 (22)	47 (21)	
Age-years	65.0 (52.0–71.5)	64.0 (51.0–70.0)	0.56
BMI-kg.m ⁻²	26.9 (23.7–30.9)	26.9 (23.6–30.9)	0.73
Unit			0.50
Surgical ICU	109 (39)	78 (35)	
Medical ICU	104 (37)	83 (37)	
Cardiac surgery ICU	66 (24)	62 (28)	
SOFA-total (/24)	11 (8–12)	11 (8–12)	0.61
Mechanical ventilation during ICU stay	256 (92)	208 (93)	0.52
ECMO during ICU stay	67 (24)	63 (28)	0.28
Length of stay-days	16.3 (8.8–31.5)	18.6 (10.1–34.1)	0.20
Anti-Xa-UI.L ⁻¹	0.29 (0.17–0.40)	0.30 (0.18–0.40)	0.83
Interval duration -hours	15.1 (11.0–19.1)	13.5 (10.3–18.8)	0.089

The features are calculated at the ICU stay level.

Data are median (Q1–Q3) for continuous variables and number (%) for categorical variables.

Statistical analyses were Wilcoxon rank sum test for continuous variables and, Pearson's Chi-squared test for categorical variables.

BMI: body mass index, ICU: intensive care unit, SOFA: sequential organ failure assessment, ECMO: extracorporeal membrane oxygenation, UFH: unfractionated heparin.

Evaluation of models without start-of-interval anti-Xa

Models trained without the start-of-interval anti-Xa feature still perform well, with an AUROC of 0.77 [0.74–0.81] and an AUPRC of 0.59 [0.54–0.63] for the random forest model in the test set.

Evaluation of the subgroup without ECMO

Models trained in the subgroup after exclusion of patients with ECMO therapy showed similar results, with the random forest model achieving the best results, yielding an AUROC of 0.80 [0.76–0.83] and an AUPRC of 0.61 [0.56–0.64] in the test set.

Discussion

The purpose of this study was to develop a machine learning-based system able to predict the anti-Xa range with readily available features. Out of several machine learning algorithms tested, random forest showed a comparatively better prediction performance. It was able to predict the anti-Xa range with an overall AUROC score of 0.80 [0.77–0.83], an overall AUPRC score of 0.61 [0.58–0.65], accuracy of 0.71 [0.68–0.73], a F1 score of 0.57 [0.53–0.61], and good calibration. These models could serve as clinical decision support systems, assisting clinicians in selecting the UFH dose most likely to achieve the targeted anti-Xa range as quickly as possible.

As expected, the UFH rate during the current interval, the UFH rate before the current interval and the start-of-interval anti-Xa were the most useful features, while less used features were gender, renal replacement therapy, and ECMO. However, models do not rely only on the start-of-interval anti-Xa, as evidenced by results achieved by models trained without this feature.

Surprisingly, the UFH bolus data was not often used by the algorithms for which feature importance was available. This may be due to the fact that UFH boli are rarely used and some bolus data may be lacking.

Heparin resistance, often caused by acquired antithrombin (AT) deficiency, is a known complication in ECMO patients.²² Since antithrombin is essential for UFH efficacy, patients with low AT levels are less responsive to standard UFH doses, making it difficult to predict the anti-Xa range in these cases. In our study, ECMO patients were excluded from the sub-analysis to ensure that the model's performance was not compromised, as our models do not incorporate antithrombin levels. Future research should aim to integrate antithrombin levels, to improve the model's ability to predict anti-Xa levels in this specific population.

Several studies have already developed machine learning models to predict aPTT in ICU settings^{15–17,23} using the publicly-available MIMIC-II or MIMIC-III databases. However, to our knowledge, our study is the first to focus on the anti-Xa assay as an outcome measure in critically ill patients. Indeed, all these studies focused on the first aPTT result after initiation of UFH therapy, and found similar accuracy to our study. Interestingly, the algorithms with the best performance differed between the studies despite using the same database. Moreover, they did not

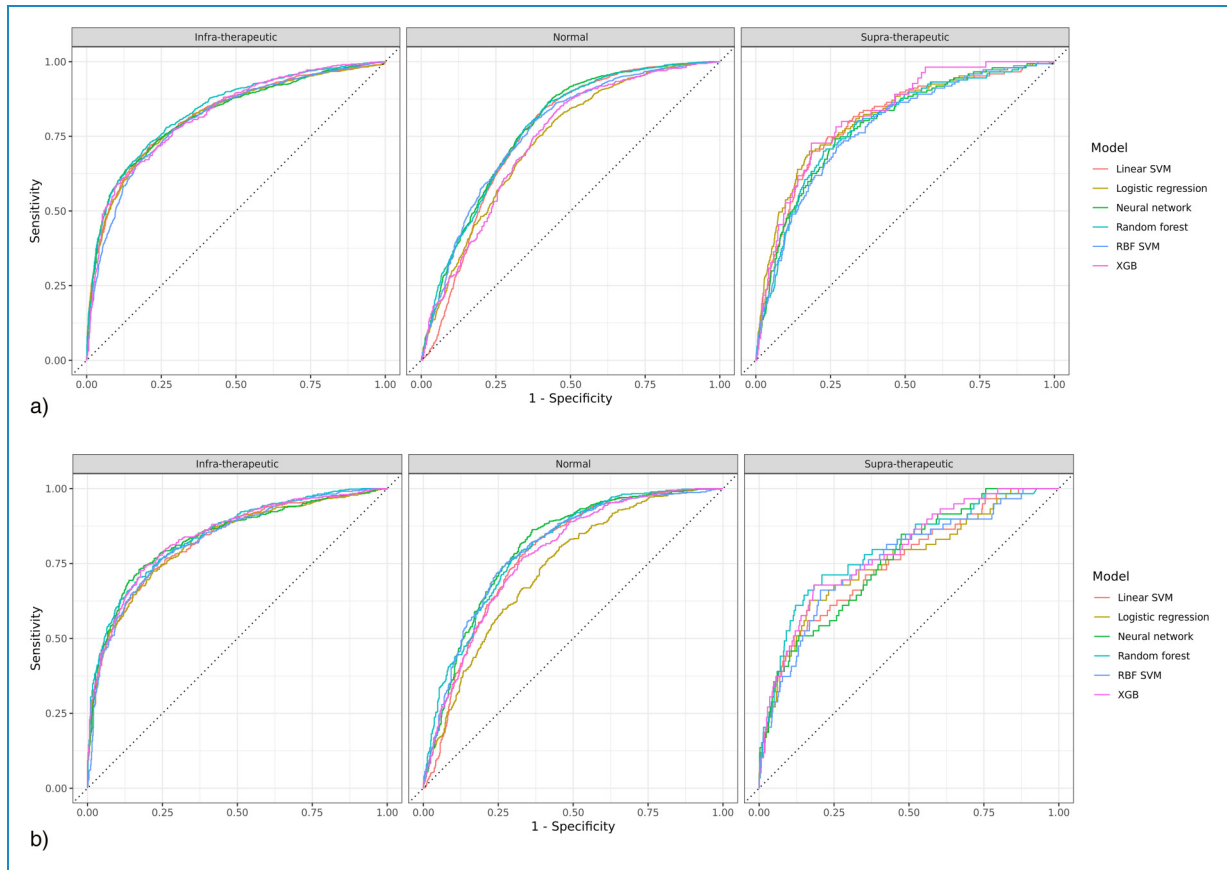


Figure 2. Receiver operating characteristic (ROC) curves of models on training (2a) and test (2b) sets. SVM, support vector machine; RBF SVM, radial basis function kernel support vector machine; XGB, extreme gradient boosting.

collect data on concomitant anticoagulant interactions, inflammation markers, interactions with ECMO or RRT. Some of these studies used reinforcement learning methods, making the results difficult to compare.

Anti-Xa is a more costly but more reliable assay than aPTT for assessing the concentration of unfractionated heparin in the blood, especially in critically ill patients, where the correlation between aPTT and anti-Xa is poor.^{24,25}

Raschke's algorithm,⁵ based solely on patient weight, has been used worldwide since 1993. However, as UFH is influenced by many parameters, especially in critically ill patients, more complex models are needed, taking into account these parameters. Thus, the accessibility of big data and the development of machine learning algorithms could be useful for building reliable models.²⁶

Nemati et al.¹⁶ developed an interesting approach using a reinforcement learning (RL) algorithm. In fact, RL is well-suited for the medication dosing problem given the sequential nature of clinical treatment, where multiple treatment decisions are performed without immediate knowledge of effectiveness. The other supervised machine learning algorithms predict an immediate result, without considering the

longer term results, which makes prediction difficult when there is no strictly one-to-one correspondence between actions and outcomes. Markov decision processes combined with RL techniques can take these long-term effects into account.

Some studies have used population pharmacokinetics/pharmacodynamics (PK/PD) models to predict hemostatic tests,^{27,28} revealing interesting results with a two-compartment model. These two studies involved patients undergoing surgery with a cardiopulmonary bypass.

Certain limitations should be acknowledged. First, because this was a single-center study, the results cannot be generalized to other hospitals. In addition, our sample size is small and may affect the accuracy of our algorithms. However, the single-center design ensured good performance with prediction models adapted to our population. To apply our model to clinical practice, predictive accuracy should be verified using a prospective external validation set and the optimal cutoff value investigated. Second, we used retrospective data. Thrombotic and hemorrhagic events were not available but would have been interesting to assess. Third, our models require the start-of-interval anti-Xa value and are therefore not suitable for predicting

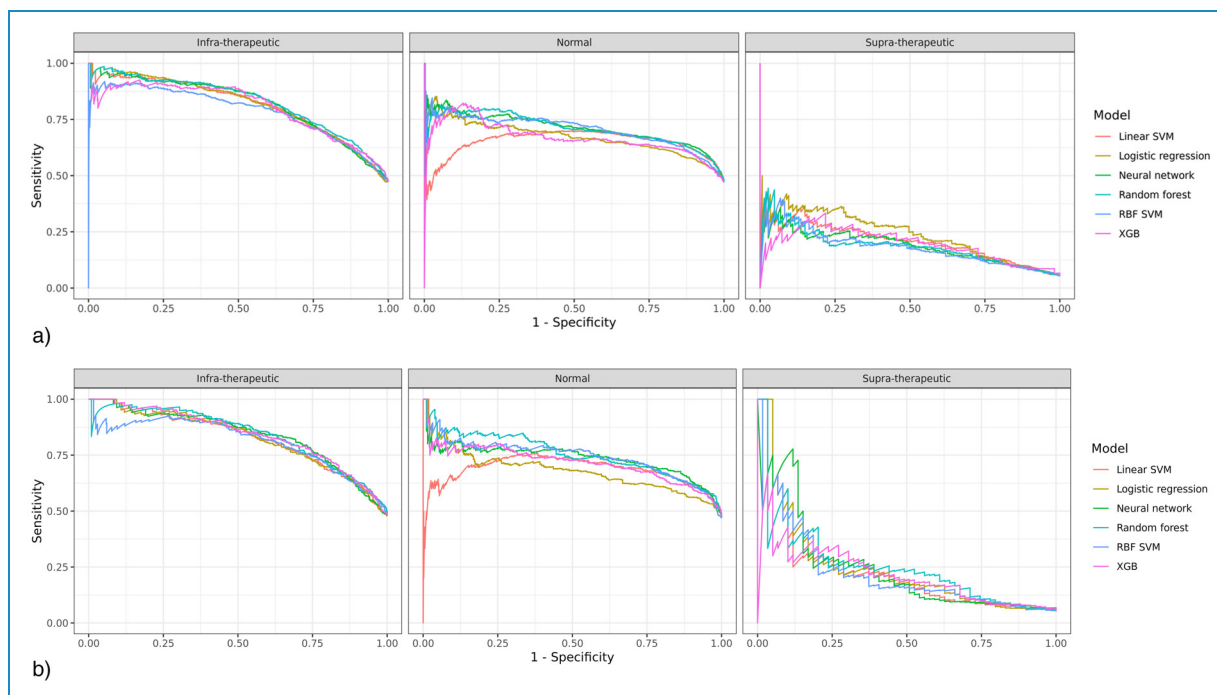


Figure 3. Precision-recall (PR) curves of models on training (3a) and test (3b) sets. SVM, support vector machine; RBF SVM, radial basis function kernel support vector machine; XGB, extreme gradient boosting.

the correct dose on initiation of UFH therapy. Finally, our dataset was imbalanced due to the relative rarity of the supra-therapeutic class, which was influenced by the small sample size. Although we applied case weights to mitigate the impact of this imbalance, the models still have limited performance in predicting supra-therapeutic outcomes.

Conclusion

In this study, the machine learning models achieved good performance in predicting anti-Xa results during UFH therapy using routinely available features. Further validation with prospective multicenter data is needed to confirm the model's generalizability and facilitate its integration into clinical practice, helping clinicians select the optimal heparin dose and reduce the risks associated with UFH therapy.

Abbreviations

aPTT	Activated partial thromboplastin time
AT	Antithrombin
AUPRC	Area under the precision-recall curve
AUROC	Area under the receiver operating characteristic curve
BMI	Body mass index
ECMO	Extracorporeal membrane oxygenation
ICU	Intensive care unit
PK	Pharmacokinetic
PK/PD	Pharmacokinetics/Pharmacodynamics
PR	Precision-Recall
RBF SVM	Radial basis function kernel support vector machine

RL	Reinforcement Learning
RRT	Renal replacement therapy
SOFA	Sequential Organ Failure Assessment
SVM	Support vector machine
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
UFH	Unfractionated heparin
XGB	eXtreme gradient boosting trees

Availability of data: The data that support the findings of this study are available from the corresponding author, BD, upon reasonable request.

Contributorship: BD conceptualized the study and participated in its design, data acquisition and analysis, literature research, and manuscript drafting. JMT, AG, and AM conceptualized the study and participated in its design, data analysis, and manuscript drafting and revision for important intellectual content. All authors read and approved the final manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics approval: This study was approved by the Rennes University Hospital Ethics Committee (no. 23.03) and written informed consent was waived from all participants.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Guarantor: As the guarantor for this manuscript, I, Boris Delange, affirm my responsibility for the integrity of the work presented. I have overseen the study design, data collection, and analysis processes to ensure accuracy and reliability. My involvement in each aspect of the research guarantees that all findings are reported honestly and transparently.

ORCID iD: Boris Delange  <https://orcid.org/0009-0002-6055-6935>

Supplemental material: Supplemental material for this article is available online.

References

1. Bjornsson TD, Wolfram KM and Kitchell BB. Heparin kinetics determined by three assay methods. *Clin Pharmacol Ther* 1982; 31: 104–113.
2. Cossette B, Pelletier ME, Carrier N, et al. Evaluation of bleeding risk in patients exposed to therapeutic unfractionated or low-molecular-weight heparin: a cohort study in the context of a quality improvement initiative. *Ann Pharmacother* 2010; 44: 994–1002.
3. Smith SB, Geske JB, Maguire JM, et al. Early anticoagulation is associated with reduced mortality for acute pulmonary embolism. *Chest* 2010; 137: 1382–1390.
4. Hull RD, Raskob GE, Brant RF, et al. Relation between the time to achieve the lower limit of the APTT therapeutic range and recurrent venous thromboembolism during heparin treatment for deep vein thrombosis. *Arch Intern Med* 1997; 157: 2562–2568.
5. Raschke R, Reilly B, Guidry J, et al. The weight-based heparin dosing nomogram compared with a standard care nomogram. *Ann Intern Med* 1993; 119: 874–881.
6. Derbalah A, Duffull S, Newall F, et al. Revisiting the pharmacology of unfractionated heparin. *Clin Pharmacokinet* 2019; 58: 1015–1028.
7. Boneu B, Caranobe C and Sie P. Pharmacokinetics of heparin and low molecular weight heparin. *Baillieres Clin Haematol* 1990; 3: 531–544.
8. White D, MacDonald S, Bull T, et al. Heparin resistance in COVID-19 patients in the intensive care unit. *J Thromb Thrombolysis* 2020; 50: 287–291.
9. George C, Barras M, Coombes J, et al. Unfractionated heparin dosing in obese patients. *Int J Clin Pharm* 2020; 42: 462–473.
10. Ling HZ, Flint J, Damgaard M, et al. Calculated plasma volume status and prognosis in chronic heart failure. *Eur J Heart Fail* 2015; 17: 35–43.
11. Kostousov V, Nguyen K, Hundalani SG, et al. The influence of free hemoglobin and bilirubin on heparin monitoring by activated partial thromboplastin time and anti-Xa assay. *Arch Pathol Lab Med* 2014; 138: 1503–1506.
12. Finley A and Greenberg C. Review article: heparin sensitivity and resistance: management during cardiopulmonary bypass. *Anesth Analg* 2013; 116: 1210–1222.
13. Rosborough TK. Monitoring unfractionated heparin therapy with antifactor Xa activity results in fewer monitoring tests and dosage changes than monitoring with the activated partial thromboplastin time. *Pharmacotherapy* 1999; 19: 760–766.
14. Guervil DJ, Rosenberg AF, Winterstein AG, et al. Activated partial thromboplastin time versus antifactor Xa heparin assay in monitoring unfractionated heparin by continuous intravenous infusion. *Ann Pharmacother* 2011; 45: 861–868.
15. Ghassemi MM, Richter SE, Eche IM, et al. A data-driven approach to optimized medication dosing: a focus on heparin. *Intensive Care Med* 2014; 40: 1332–1339.
16. Nemati S, Ghassemi MM and Clifford GD. Optimal medication dosing from suboptimal clinical examples: a deep reinforcement learning approach. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016: pp.2978–2981.
17. Su L, Liu C, Li D, et al. Toward optimal heparin dosing by comparing multiple machine learning methods: retrospective study. *JMIR Med Inform* 2020; 8: e17648.
18. Madec J, Bouzillé G, Riou C, et al. eHOP clinical data warehouse: from a prototype to the creation of an inter-regional clinical data centers network. *Stud Health Technol Inform* 2019; 264: 1536–1537.
19. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; 162: 55–63.
20. Abe D, Inaji M, Hase T, et al. A prehospital triage system to detect traumatic intracranial hemorrhage using machine learning algorithms. *JAMA Netw Open* 2022; 5: e2216393.
21. Hand DJ and Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Mach Learn* 2001; 45: 171–186.
22. Raghunathan V, Liu P, Kohs TCL, et al. Heparin resistance is common in patients undergoing extracorporeal membrane oxygenation but is not associated with worse clinical outcomes. *ASAIO J Am Soc Artif Intern Organs* 2021; 67: 899–906.
23. Lin R, Stanley MD, Ghassemi MM, et al. A deep deterministic policy gradient approach to medication dosing and surveillance in the ICU. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018. pp.4927–4931.
24. Aarab R, Es Jv, Pont Ad, et al. Monitoring of unfractionated heparin in critically ill patients. *Neth J Med* 2013; 71: 466–471.
25. Ratano D, Alberio L, Delodder F, et al. Agreement between activated partial thromboplastin time and anti-Xa activity in critically ill patients receiving therapeutic unfractionated heparin. *Thromb Res* 2019; 175: 53–58.
26. Sidey-Gibbons JAM and Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol* 2019; 19: 64.
27. Delavenne X, Ollier E, Chollet S, et al. Pharmacokinetic/pharmacodynamic model for unfractionated heparin dosing during cardiopulmonary bypass. *Br J Anaesth* 2017; 118: 705–712.
28. Jia Z, Tian G, Ren Y, et al. Pharmacokinetic model of unfractionated heparin during and after cardiopulmonary bypass in cardiac surgery. *J Transl Med* 2015; 13: 45.