

Explainable Machine Learning Model for Predicting Filter Lifespan in Patients on Continuous Renal Replacement Therapy

Sheng-kai Pan

Zhejiang Chinese Medical University

Cheng Zheng

Zhejiang Chinese Medical University

Xiang-hong Yang

yangxianghong@hmc.edu.cn

Zhejiang Provincial People's Hospital

Research Article

Keywords: Continuous renal replacement therapy, Filter lifespan, Machine learning, Random forest model, SHAP ana, Predictive model

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10894003/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Continuous renal replacement therapy (CRRT) is a key modality for sustaining patients' vital signs. The filter is critical for achieving therapeutic efficacy, and its lifespan is influenced by a complex interplay of multiple factors, including patient coagulation status, anticoagulation regimen, vascular access, and operational parameters. This study aimed to develop an interpretable machine learning model based on clinical data to predict early filter failure risk in CRRT, and to employ SHAP analysis to elucidate the influence of various factors on filter lifespan.

Methods

This retrospective study included 854 patients, from whom clinical data were collected. We evaluated six machine learning models for predicting factors affecting filter lifespan, and assessed their performance using the area under the receiver operating characteristic curve (AUC), average precision (AP), confusion matrix, and other metrics. SHAP analysis was applied to determine feature importance and to interpret the machine learning models.

Results

Among the six evaluated models, the random forest (RF) model achieved the highest performance in the validation set, with an AUC of 0.805 (95% CI: 0.744–0.863), an AP of 0.642 (95% CI: 0.523–0.769), and an F1 score of 0.633 (95% CI: 0.572–0.690). SHAP analysis revealed that anticoagulation modality (particularly heparin-free and citrate anticoagulation), puncture site, age, serum potassium, international normalized ratio (INR), and estimated glomerular filtration rate (eGFR) were the most important factors influencing filter lifespan. At the individual patient level, SHAP force plots provided in-depth, interpretable visualizations of the predictions generated by our random forest model.

Conclusions

The Random forest model identified as the most effective for early filter failure risk prediction in patients receiving CRRT. By incorporating SHAP analysis, clinicians can clearly understand the key determinants of filter lifespan and their directions of effect, thereby enabling proactive optimization of anticoagulation strategies and vascular access management in high-risk patients to extend filter longevity and reduce unplanned treatment interruptions.

Trial registration

Not applicable.

Introduction

Continuous renal replacement therapy (CRRT), as a core life-sustaining modality in the intensive care unit (ICU) ^[1] not only serves as renal replacement therapy but also contributes to maintaining internal environment homeostasis, and is primarily used to treat acute kidney injury. By removing cytotoxins, metabolic wastes, and inflammatory factors through diffusion or convection, CRRT improves tissue oxygen metabolism and sustains internal environment stability ^[2].

The filter is the core component of CRRT, and its lifespan is a major determinant of treatment efficacy and patient prognosis [3]. As the site where fluid and/or solute imbalances are corrected, the filter is also the most susceptible to clotting. In clinical practice, approximately 75.7% of patients receiving CRRT experience premature treatment termination due to unanticipated filter clotting [4], with about 33.9% of events occurring within 24 hours and 30.7% between 24 and 48 hours [5]. Moreover, various other factors—including patient-specific characteristics, anticoagulant selection, catheter-related issues, and CRRT operational parameters—can also shorten filter lifespan and consequently compromise therapeutic outcomes. Therefore, accurately identifying patients at high risk of early filter failure and implementing effective interventions in the context of complex, multifactorial clinical scenarios, so as to extend filter longevity and ensure CRRT efficacy, has emerged as a pivotal clinical challenge.

With the ongoing advancement of artificial intelligence, the application of machine learning (ML) has become increasingly prevalent in healthcare. Clinically, ML enhances decision-making through early warning, diagnostic facilitation, personalized treatment, and assessment of patient responses to therapy. Tseng et al. developed an ML-based model to predict acute kidney injury (AKI) after cardiac surgery, demonstrating that their XGBoost model significantly outperformed conventional logistic regression [6]. Kang et al. reported that prediction models built using ML algorithms—including random forest and support vector machine—based on data collected at CRRT initiation in AKI patients achieved higher accuracy than traditional Acute Physiology and Chronic Health Evaluation II (APACHE II) and Sequential Organ Failure Assessment (SOFA) scores [7].

Although ML has made remarkable and sustained progress over the past decade, ML models are often regarded as opaque "black boxes." Recent advances in explainable artificial intelligence (XAI) have facilitated the interpretability of these models. Currently, ML research specifically focused on predicting CRRT filter lifespan remains scarce. We therefore hypothesized that ML models could enable clinicians to predict filter lifespan using routine clinical data, and that the application of XAI approaches could provide interpretable predictions and deeper insights into filter longevity. Accordingly, this study aimed to develop an interpretable ML model using routinely available clinical data to identify patients at high risk of early filter replacement, and to provide individualized explanations via SHAP to support decision-making in anticoagulation management.

Materials and methods

1. Study population

This retrospective study was designed to develop and validate machine learning (ML) models for predicting filter lifespan in patients receiving CRRT. Data were derived from patients treated in the ICU of Zhejiang Provincial People's Hospital between 2021 and 2024. Inclusion criteria were: (1) receipt of CRRT; and (2) age ≥ 18 years. Exclusion criteria were: (1) treatment duration < 24 hours; (2) missing data exceeding 30% for key variables; and (3) filter replacement due to non-clotting reasons (e.g., patient transport or surgical interruption). A total of 854 patients were ultimately included. This study was approved by the hospital ethics committee, and the requirement for informed consent was waived.

2. Data processing

Feature selection was performed based on domain expertise, incorporating variables with potential discriminative ability for predicting filter lifespan [8][9][10]. The selected features comprised: (1) demographic characteristics: age and sex; (2) laboratory tests: lactate, prothrombin time, international normalized ratio (INR), urea, white blood cell count, albumin, bicarbonate, fibrinogen, platelet count, hemoglobin, high-sensitivity C-reactive protein, activated partial thromboplastin time, pH, serum sodium, serum potassium, serum magnesium, procalcitonin, creatinine, and estimated

glomerular filtration rate (eGFR); (3) CRRT-related data: anticoagulation modality (heparin, heparin-free, nafamostat, and citrate), arterial pressure, venous pressure, transmembrane pressure, dehydration volume, blood flow rate, ultrafiltration rate, dialysate flow rate, machine model (Gambro, Fresenius, Baxter, or others), anticoagulation mode (CVVH, CVVHD, CVVHDF), and puncture site (left femoral vein, left internal jugular vein, right femoral vein, right internal jugular vein, or ECMO circuit); and (4) comorbidities: hypertension, diabetes, chronic kidney disease, arrhythmia, heart failure, chronic obstructive pulmonary disease (COPD), cirrhosis, and infection. All laboratory values were obtained from the measurement closest to CRRT initiation within the preceding 24 hours. To standardize the data, features were rescaling using z-scores.

Given the presence of missing data, variables with a missing proportion exceeding 30% were excluded from subsequent analyses to mitigate their impact on model stability, while variables with missing proportions $\leq 30\%$ were retained. For missing values within the retained variables, multiple imputation was applied.

3. Outcome variable

In accordance with guidelines from the International Society of Nephrology (ISN) and the Acute Dialysis Quality Initiative (ADQI), the effective lifespan of the CRRT filter is a critical indicator of treatment continuity and adequacy^[11]^[12]. In this study, the actual service life of the first filter used during each CRRT session was defined as the observation endpoint and was transformed into a binary outcome variable based on the clinically acceptable minimum effective treatment duration threshold. Specifically, a single filter lifespan exceeding 60 hours was defined as favorable (coded as 1), while a lifespan ≤ 60 hours was defined as unfavorable (coded as 0). The 60-hour cutoff was selected based on previously reported average filter replacement intervals in CRRT^[12] and a comprehensive consideration of tolerance to treatment interruption in clinical practice. A lifespan beyond this threshold indicates satisfactory anticoagulant efficacy and hemodynamic compatibility, whereas a shorter duration suggests an elevated risk of early clotting or filter dysfunction.

4. Machine learning model development

Data from patients who received CRRT at the ICU of Zhejiang Provincial People's Hospital between 2021 and 2024 were selected and stratified random sampling was performed to divide the data into a training set and an internal validation set at a 7:3 ratio. The training set was used for feature selection, model training, and hyperparameter optimization, while the internal validation set was used to evaluate the discriminative ability, calibration, and classification performance of the models on unseen data.

To reduce the risk of overfitting due to high-dimensional features and to identify candidate predictors associated with the outcome, LASSO binary logistic regression was applied to the training set for feature selection prior to subsequent machine learning modeling. We constructed and compared six machine learning classifiers: (1) logistic regression (LR); (2) random forest (RF); (3) gradient boosting classifier (GBC); (4) adaboost classifier (ABC); (5) extreme gradient boosting (XGB); and (6) light gradient boosting machine (LGB). The dataset was partitioned into training and internal validation sets at a 7:3 ratio. Given the frequent class imbalance between positive and negative samples in binary medical data, class imbalance handling was performed during model training. The Bootstrap method with 500 repetitions was used to derive the area under the receiver operating characteristic curve (AUC) and average precision (AP). Based on the optimal threshold determined from the training set, confusion matrices were generated for both the training and internal validation sets, and sensitivity, specificity, positive predictive value, negative predictive value, accuracy, F1 score, precision, and recall were subsequently calculated. AUC, AP, and F1 score were the primary metrics for model performance comparison^{[13][14][15][16]}. Hyperparameter optimization for the six ML models was performed

using stratified 5-fold cross-validation. To further mitigate the risk of overfitting associated with small-sample modeling, an overfit guard strategy was incorporated into the scripting pipeline.

To interpret the model predictions, SHAP analysis was performed separately for the RF, LR, GBC, ABC, XGB, and LGB models. For tree-based models, TreeExplainer was preferentially used to compute SHAP values; for linear models, an interpreter appropriate for the model type was applied; for other models, permutation-based interpretation methods were used based on model compatibility. SHAP feature importance plots displayed the global importance of each feature, with a larger mean absolute SHAP value indicating greater contribution to the model predictions. SHAP summary plots illustrated the direction and magnitude of influence of each feature on the predictions. In the SHAP summary plots, each point represents the SHAP value of a given feature for an individual patient, with colors ranging from red to blue indicating high to low feature values.

5. Statistical methods

Categorical variables were presented as frequencies and percentages, while continuous variables were expressed as mean $\pm$ standard deviation or median (interquartile range) depending on their distribution. Comparisons between groups were performed using appropriate statistical tests based on variable types and distributions. Data preprocessing, missing value imputation, baseline characteristic comparisons, and LASSO feature selection were conducted in the R software environment, primarily using the readxl, openxlsx, mice, caret, CBCgrps, glmnet, and ggplot2 packages. Machine learning model construction, performance evaluation, plotting, and SHAP analysis were performed in the Python environment, primarily using the pandas, numpy, scikit-learn, xgboost, lightgbm, matplotlib, joblib, and shap packages. Random seeds were set for all stochastic processes to enhance the reproducibility of the study results.

Results

1. Baseline characteristics

A total of 854 patients receiving CRRT were included in this study, of whom 228 (26.7%) had a favorable filter lifespan (> 60 h) and 626 (73.3%) had an unfavorable filter lifespan (≤ 60 h). In the training set, 160 patients (27%) exhibited a favorable lifespan, while 68 patients (27%) did so in the internal validation set.

Patient characteristics are summarised in Table 1. Sex, comorbidities (hypertension, diabetes, chronic kidney disease, arrhythmia, heart failure, COPD, cirrhosis, and infection), and most laboratory parameters (INR, haemoglobin, lactate, CRP, etc.) showed no statistically significant differences between groups (all $P > 0.05$). In both the training and internal validation sets, anticoagulation modality and eGFR were significantly associated with filter lifespan; however, no significant differences were observed between the two groups regarding anticoagulation modality ($P = 0.627$) or eGFR (40.39 vs. 38.38, $P = 0.479$). Patients with lower eGFR tended to have a longer filter lifespan.

Table 1
Baseline characteristics of the training and internal validation sets

Characteristic	Training set		Internal validation set			
	Unfavorable (n = 439)	Favorable (n = 160)	P value	Unfavorable (n = 187)	Favorable (n = 68)	P value
Sex			0.74			0.96
Male	299 (68)	106 (66)		124 (66)	46 (68)	
Female	140 (32)	54 (34)		63 (34)	22 (32)	
Anticoagulation modality			< 0.001			< 0.001
Heparin	221 (50)	73 (46)		96 (51)	30 (44)	
Heparin-free	146 (33)	26 (16)		61 (33)	10 (15)	
Nafamostat	32 (7)	15 (9)		12 (6)	3 (4)	
Citrate	40 (9)	46 (29)		18 (10)	25 (37)	
Machine model			0.309			0.088
Gambro	190 (43)	77 (48)		78 (42)	25 (37)	
Fresenius	184 (42)	67 (42)		80 (43)	37 (54)	
Baxter	61 (14)	14 (9)		27 (14)	4 (6)	
Others	4 (1)	2 (1)		2 (1)	2 (3)	
CRRT modality			0.081			0.021
CVVH	44 (10)	7 (4)		27 (14)	2 (3)	
CVVHD	185 (42)	75 (47)		75 (40)	36 (53)	
CVVHDF	210 (48)	78 (49)		85 (45)	30 (44)	
Puncture site			0.075			0.217
Left femoral vein	139 (32)	35 (22)		41 (22)	18 (26)	
Left internal jugular vein	4 (1)	0 (0)		1 (1)	1 (1)	
Right femoral vein	262 (60)	109 (68)		128 (68)	43 (63)	
Right internal jugular vein	12 (3)	8 (5)		6 (3)	5 (7)	
ECMO	22 (5)	8 (5)		11 (6)	1 (1)	
Hypertension	244 (56)	91 (57)	0.85	114 (61)	48 (71)	0.206
Diabetes	142 (32)	55 (34)	0.712	58 (31)	25 (37)	0.474
Chronic kidney disease	164 (37)	63 (39)	0.722	74 (40)	33 (49)	0.255
Arrhythmia	110 (25)	35 (22)	0.486	57 (30)	17 (25)	0.486
Heart failure	142 (32)	40 (25)	0.103	56 (30)	15 (22)	0.278

Characteristic	Training set			Internal validation set		
	Unfavorable (n = 439)	Favorable (n = 160)	P value	Unfavorable (n = 187)	Favorable (n = 68)	P value
COPD	23 (5)	9 (6)	1	15 (8)	2 (3)	0.254
Cirrhosis	17 (4)	6 (4)	1	7 (4)	0 (0)	0.195
Infection	312 (71)	108 (68)	0.457	140 (75)	47 (69)	0.449
Age (years)	70 (59, 78)	72 (61, 82)	0.112	71 (59, 80)	68 (60.75, 80.25)	0.907
Lactate (mmol/L)	3.1 (1.7, 6.9)	2.7 (1.6, 6.15)	0.437	3 (1.75, 6.9)	2.25 (1.58, 5.54)	0.124
PT (s)	15 (13, 19.8)	15.35 (13.3, 18.5)	0.863	14.5 (12.9, 17.05)	14.3 (13.23, 15.53)	0.491
INR	1.37 (1.19, 1.77)	1.39 (1.19, 1.7)	0.901	1.32 (1.19, 1.6)	1.29 (1.23, 1.45)	0.528
BUN (mmol/L)	15.29 (9.39, 24.41)	17.41 (9.7, 24.16)	0.221	13.86 (8.29, 20.81)	14.16 (9.98, 21.68)	0.352
WBC (×10⁹/L)	11.18 (7.08, 16.06)	12.59 (8.88, 17.4)	0.004	12.03 (7.44, 16.91)	12.35 (9.02, 17.8)	0.434
ALB (g/L)	30.5 ± 5.26	31.19 ± 5.27	0.154	30.83 ± 5.12	28.3 ± 6.06	0.003
Bicarbonate (mmol/L)	17.9 (13.7, 21.85)	18.05 (14.8, 21.5)	0.622	18.5 (13.9, 21.9)	18.5 (14.47, 20.8)	0.964
Fibrinogen (g/L)	3.5 (2.02, 4.78)	3.63 (2.33, 4.59)	0.48	3.58 (2.29, 4.74)	3.41 (2.6, 4.78)	0.509
Platelet (×10⁹/L)	130 (76, 197)	130.5 (77.75, 199)	0.786	148 (87, 206)	127.5 (71, 195.75)	0.447
Hb (g/L)	91 (70.5, 111)	92 (77, 109)	0.238	85 (71, 113.5)	87 (72, 116)	0.87
CRP (mg/L)	74.1 (25, 146)	70.8 (31.1, 143.8)	0.888	86.2 (30.25, 147.4)	58.55 (23.07, 138.85)	0.322
APTT (s)	33.8 (28.8, 42.95)	33.8 (28.58, 43.47)	0.889	31.2 (27.3, 40.6)	31.4 (27.9, 37.6)	0.589
PH	7.33 (7.24, 7.4)	7.33 (7.23, 7.38)	0.761	7.32 (7.2, 7.38)	7.33 (7.27, 7.42)	0.161
Na⁺ (mmol/L)	139 (134.4, 148.7)	139 (134, 148)	0.663	138.9 (134.1, 146)	139.8 (135, 148.5)	0.29
K⁺ (mmol/L)	4 (3.39, 5.3)	4.14 (3.4, 5.24)	0.71	3.9 (3.48, 5.04)	3.92 (3.38, 5.4)	0.971
Mg²⁺ (mmol/L)	0.87 (0.77, 1)	0.86 (0.78, 1)	0.79	0.87 (0.78, 1)	0.88 (0.77, 1.03)	0.784
PCT (ng/mL)	1.43 (0.44, 6.7)	2.23 (0.62, 8.72)	0.134	1.52 (0.42, 5.54)	1.39 (0.57, 4.6)	0.885

Characteristic	Training set			Internal validation set		
	Unfavorable (n = 439)	Favorable (n = 160)	P value	Unfavorable (n = 187)	Favorable (n = 68)	P value
Cr (μmol/L)	172.6 (105.15, 347.2)	218.8 (122.48, 351.3)	0.062	183 (100.9, 347.15)	263.5 (124, 441.07)	0.035
eGFR(mL/min/1.73 m ²)	44.61 (19.94, 89.06)	32.01 (16.74, 68.89)	0.014	42.77 (18.75, 95.32)	22.02 (13.32, 56.12)	0.003
Arterial pressure (mmHg)	-20 (-45, -20)	-20 (-40, -20)	0.747	-25 (-40.5, -18.5)	-22.5 (-40, -20)	0.736
Venous pressure (mmHg)	40 (30, 60)	31.5 (25, 50)	0.081	39 (30, 60)	32.5 (20, 54.25)	0.106
Transmembrane pressure (mmHg)	22 (10, 42)	27 (10, 41.25)	0.897	30 (10, 42.5)	20 (10, 37)	0.032
Dehydration volume (mL)	150 (90, 200)	150 (99, 200)	0.726	152 (108, 196.5)	152.5 (97.75, 211.68)	0.781
Blood flow rate (mL/min)	100 (100, 150)	100 (100, 150)	0.322	100 (100, 150)	100 (100, 150)	0.094
Ultrafiltration rate (mL/h)	200 (150, 220)	200 (167.5, 220)	0.43	200 (170, 220)	200 (165, 220)	0.748
Dialysate flow rate (mL/h)	2000 (1000, 2000)	2000 (1000, 2000)	0.042	2000 (1000, 2000)	2000 (1000, 2000)	0.207

Data are presented as n (%) for categorical variables, mean ± SD for normally distributed continuous variables (Albumin), and median (interquartile range) for non-normally distributed continuous variables. Abbreviations: PT, prothrombin time; INR, international normalized ratio; BUN, blood urea nitrogen; WBC, white blood cell count; CRP, C-reactive protein; PCT, procalcitonin; Cr, creatinine; APTT, activated partial thromboplastin time; eGFR, estimated glomerular filtration rate; CVVH, continuous venovenous hemofiltration; CVVHD, continuous venovenous hemodialysis; CVVHDF, continuous venovenous hemodiafiltration; COPD, chronic obstructive pulmonary disease.

2. Model evaluation and interpretation

Among the six developed models, the RF model achieved the highest AUC in the validation set, at 0.805 (95% CI: 0.744–0.863), followed by the XGB model, with an AUC of 0.794 (95% CI: 0.730–0.858) (Fig. 1A). To account for class imbalance in the data, we additionally plotted precision–recall (PR) curves and calculated the average precision (AP) as a complementary metric for model performance evaluation under imbalanced conditions. In the validation set, the RF model also exhibited the highest AP, at 0.642 (95% CI: 0.523–0.769) (Fig. 1B), followed by the XGB model, with an AP of 0.640 (95% CI: 0.517–0.756) (Fig. 2).

The ROC and PR curves described above validated the effectiveness of the models from a global ranking perspective. However, in real-world deployment scenarios, a single definitive cut-off threshold must be selected for the final algorithm. To further elucidate the detailed classification performance at this threshold and to quantify the differential impact of misclassification, we introduced confusion matrices for fine-grained dissection of the model predictions. Based on these matrices, we derived core evaluation metrics, including precision, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 score^{[15][16]}.

As shown in Fig. 2 and Table 2, the RF model achieved the highest precision and F1 score, at 0.773 (95% CI: 0.717–0.820) and 0.633 (95% CI: 0.572–0.690), respectively. The LGB model demonstrated the highest sensitivity, at 0.824 (95% CI: 0.772–0.865), while the XGB model exhibited the highest specificity, at 0.802 (95% CI: 0.749–0.846). Overall, the RF model not only achieved the highest AUC and AP but also attained the best F1 score among all six classifiers.

Based on the comprehensive evaluation of key metrics, including ROC-AUC, PR-AP, and precision and F1 score derived from confusion matrices, the RF model demonstrated the best overall performance among the six classifiers. It achieved the highest AUC (0.805) and AP (0.642), as well as the highest F1 score (0.633), indicating an optimal trade-off between precision and sensitivity. Although the LGB model showed a slight advantage in sensitivity and the XGB model exhibited superior specificity, the RF model proved to be the most robust in balancing the identification of both positive and negative classes while controlling the cost of misclassification. Therefore, the RF model was selected as the final optimal classifier for further SHAP analysis.

Table 2
Metrics on the validation set (including precision, sensitivity, specificity, PPV, NPV, and F1 score)

Model	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1	Optimal_Threshold
ABC	0.781 (0.716–0.846)	0.745 (0.688–0.795)	0.765 (0.709–0.813)	0.738 (0.681–0.788)	0.515 (0.454–0.576)	0.896 (0.853–0.928)	0.615 (0.554–0.673)	0.375475
GBC	0.792 (0.731–0.854)	0.698 (0.639–0.751)	0.735 (0.678–0.786)	0.684 (0.625–0.738)	0.459 (0.399–0.520)	0.877 (0.831–0.912)	0.565 (0.504–0.624)	0.258948
LGB	0.791 (0.727–0.852)	0.620 (0.559–0.677)	0.824 (0.772–0.865)	0.545 (0.484–0.605)	0.397 (0.339–0.458)	0.895 (0.851–0.927)	0.536 (0.475–0.596)	0.442174
LR	0.756 (0.690–0.816)	0.616 (0.555–0.673)	0.765 (0.709–0.813)	0.561 (0.500–0.621)	0.388 (0.330–0.449)	0.868 (0.821–0.904)	0.515 (0.454–0.576)	0.437486
RF	0.805 (0.744–0.863)	0.773 (0.717–0.820)	0.735 (0.678–0.786)	0.786 (0.732–0.832)	0.556 (0.494–0.615)	0.891 (0.847–0.924)	0.633 (0.572–0.690)	0.517086
XGB	0.794 (0.730–0.858)	0.769 (0.713–0.816)	0.676 (0.617–0.731)	0.802 (0.749–0.846)	0.554 (0.493–0.614)	0.872 (0.826–0.908)	0.609 (0.548–0.667)	0.545862

The SHAP feature importance of the RF model is presented in Fig. 3A. The features are ranked in descending order by their mean absolute SHAP values, reflecting their relative influence on model predictions. The seven most important features were heparin-free anticoagulation, citrate anticoagulation, INR, serum potassium, eGFR, right internal jugular vein as puncture site, and age.

The SHAP summary plot further illustrated the directional impact of each feature on model predictions (Fig. 3B). According to the prediction model, higher SHAP values were associated with a greater likelihood of favorable filter lifespan. For instance, the use of heparin-free anticoagulation was associated with a higher probability of filter lifespan ≤ 60 hours, whereas citrate anticoagulation was more likely to extend filter lifespan beyond 60 hours.

In addition, SHAP dependence plots were generated to visualize the marginal effects of individual features on RF predictions (Fig. 4). Higher INR, higher serum potassium, and older age (increasing values along the x-axis) were associated with higher SHAP values (increasing values along the y-axis), indicating a greater likelihood of favorable

filter lifespan. Conversely, lower eGFR was associated with higher SHAP values, suggesting that patients with more severely reduced renal function tended to have prolonged filter longevity.

3. Patient-level interpretation of the machine learning model

To evaluate the contribution of individual features to the predictions made by the RF model, we applied SHAP analysis to interpret the individualized predictions for two representative patients. Patient A (Fig. 5 A) was an 82-year-old patient whose actual filter lifespan exceeded 60 hours, with a model-predicted probability of favorable filter lifespan of 75.6%. This patient received citrate anticoagulation, had the catheter placed in the right internal jugular vein, and presented with an INR of 1.85 and a serum potassium level of 5.71 mmol/L. The SHAP analysis revealed that citrate anticoagulation, right internal jugular vein access, elevated INR, and elevated serum potassium were the major factors contributing to a favorable filter lifespan.

In contrast, Patient B (Fig. 5 B) was a 69-year-old patient receiving heparin-free anticoagulation, also with right internal jugular vein access, but with an INR of 0.996 and a serum potassium level of 3.39 mmol/L. The SHAP analysis indicated that heparin-free anticoagulation, low INR, and low serum potassium were the predominant contributors to an unfavorable filter lifespan.

Discussion

Based on multidimensional clinical data from 854 critically ill patients receiving CRRT, we developed and compared six machine learning models for predicting early filter failure risk. Our results demonstrated that the random forest (RF) model exhibited superior discriminative ability (AUC = 0.805, AP = 0.642) and classification performance (F1 score = 0.633) on the internal validation set, significantly outperforming the other algorithms. Through SHAP analysis, we further elucidated the key features influencing filter lifespan and their directions of effect, and achieved visualization and interpretability of predictions at the individual patient level. This study fills a gap in the application of machine learning to CRRT filter lifespan prediction and provides a novel tool with both high accuracy and high transparency for clinical anticoagulation management and vascular access decision-making.

SHAP analysis not only enhanced the interpretability of the model but also yielded core findings with direct clinical practice implications. First, anticoagulation modality was the most important feature affecting filter lifespan. Heparin-free anticoagulation was closely associated with shortened filter lifespan, whereas citrate anticoagulation significantly favored extended filter duration. This finding is consistent with existing clinical observations: heparin-free anticoagulation lacks an effective anticoagulation target, leading to a markedly elevated risk of filter clotting^{[17][18]}. A randomized controlled trial^[19] comparing regional citrate anticoagulation with systemic heparinization reported that regional citrate prolonged filter lifespan by approximately 11 hours. Additionally, puncture site also exerted a measurable impact on filter lifespan; SHAP analysis suggested that the right internal jugular vein was more likely to be associated with prolonged filter longevity. The KDIGO guidelines recommend the right internal jugular vein as the preferred site for catheter insertion, followed by the femoral vein, left internal jugular vein, and subclavian vein^{[12][23]}. Catheters placed in the right internal jugular vein follow a straighter path into the superior vena cava with minimal vessel wall contact, thereby reducing the risk of stenosis and thrombotic complications.

Second, a higher INR was found to be associated with prolonged filter lifespan to some extent. INR is a measure of coagulation time; elevated INR indicates a relatively "diluted" anticoagulant state with attenuated coagulation function^[20]. One study noted that low platelet counts in critically ill patients receiving CRRT served as a protective factor for extended filter lifespan^[21], which aligns with the protective effect of INR on filter longevity through a similar mechanism. Furthermore, elevated serum potassium was identified as a positive predictor of favorable filter lifespan.

An early in vitro experimental study suggested that K^+ can inhibit thromboplastin, thereby slowing coagulation [22]. Moreover, as both K^+ and Ca^{2+} are cations, high concentrations of K^+ may competitively interfere with the normal function of calcium ions in the coagulation cascade, further retarding the clotting process. In addition, lower eGFR indicates impaired renal clearance of metabolic waste products, which may affect coagulation function and delay filter clotting.

However, it should be noted that the proposed effects of INR, hyperkalemia, and low eGFR on filter lifespan currently remain at the level of theoretical inference based on coagulation physiology, without strong direct clinical evidence. The factors influencing CRRT filter lifespan are numerous and complex; rather than indicating a genuine protective effect per se, these associations may more likely reflect that patients with more pronounced coagulation disorders or metabolic derangements tend to receive more aggressive anticoagulation strategies or closer monitoring, thereby indirectly prolonging filter duration. Moreover, SHAP analysis suggested that the relative importance of these factors was low, indicating that they should serve only as reference points in clinical decision-making rather than as definitive therapeutic targets.

Several limitations of this study should be acknowledged when interpreting and generalizing the findings. First, this was a single-center retrospective study with data derived from the ICU of Zhejiang Provincial People's Hospital; patient populations, CRRT operational protocols, and anticoagulation strategies may differ across centers, and the external generalizability of the model requires further validation through multi-center prospective studies. Second, we used a binary cutoff of 60 hours to define favorable versus unfavorable filter lifespan. Although this threshold was informed by the literature and clinical practice, institutional variations in tolerance for filter replacement may affect the model's adaptability to different clinical settings. Third, this study only predicted the lifespan of the first filter used during each CRRT session and did not account for sequential dependencies when patients received multiple filters or for time-varying covariates. Future studies could incorporate dynamic survival analysis models to provide more refined time-to-event risk predictions.

Conclusions

In this study, we successfully developed and validated an interpretable machine learning model based on the random forest algorithm for predicting early filter failure risk in patients receiving CRRT. The model demonstrated satisfactory performance in internal validation (AUC = 0.805, AP = 0.642, F1 = 0.633), and the application of SHAP analysis enabled transparent interpretation at both the global and individual patient levels. Key risk and protective factors identified included anticoagulation modality—particularly heparin-free versus citrate anticoagulation—along with INR, serum potassium, eGFR, and puncture site, with the right internal jugular vein emerging as a favorable access site. This model holds promise as a practical decision-support tool in clinical CRRT management, facilitating early identification of high-risk patients and optimization of anticoagulation and catheterization strategies, thereby extending filter lifespan and improving patient outcomes.

Declarations

Ethics approval and consent to participate

This study complies with the Declaration of Helsinki, and the Ethical Committee of Zhejiang Provincial People's Hospital approved this study and approved the exemption of participant informed consent.

Competing interests

The authors declare no competing interests.

Funding

This research did not receive funding.

Author Contribution

S.P. performed data analysis, manuscript writing and revision of the manuscript, X. Y. and C. Z. performed the data collection as well as data interpretation.

Acknowledgements

All authors read the final version, and there is no conflict of interest.

Data Availability

Data used in this study are available from first author.

References

1. TEIXEIRA JP, MAYER K P, GRIFFIN B R, et al. Intensive care unit-acquired weakness in patients with acute kidney injury:A contemporary review[J]. Am J Kidney Dis. 2023;81(3):336–51.
2. Gaudry S, Grolleau F, Barbar S, et al. Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU studies[J]. Crit Care. 2022;26(1):93. 10.1186/s13054-022-03955-9.
3. Zhou ZF, Liu C, Yang YY, et al. Anticoagulation options for continuous renal replacement therapy in critically ill patients: a systematic review and network meta-analysis of randomized controlled trials[J]. Crit Care. 2023;27(1):222. 10.1186/s13054-023-04519-1.
4. Li PY, Zhang L, Lin L et al. Effect of dynamic circuit pressures monitoring on the lifespan of extracorporeal circuit and the efficiency of solute removal during continuous renal replacement therapy[J]. Front Med (Lausanne), 2021, 8: 621921. 10.3389/fmed. 2021. 621921.
5. Yoshimoto K, Matsuura R, Komaru Y, et al. Solute clearance evaluation and filter clotting prediction in continuous renal replacement therapy[J]. J Clin Med. 2023;12(24):7703. 10.3390/jcm12247703.
6. TSENG P Y, CHEN Y T, WANG C H, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning [J]. Crit Care. 2020;24(1):478.
7. KANG M W, KIM J, KIM D K, et al. Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy [J]. Crit Care. 2020;24(1):42.
8. Karkar A, Ronco C. Prescription of CRRT: a pathway to optimize therapy. Ann Intensive Care. 2020;10(1):32. 10.1186/s13613-020-0648-y. . PMID: 32144519; PMCID: PMC7060300.
9. Wu X, Ye J, Sun M, Wang Z, Chen Q, Zhu J. [Relationship between the timing of initiation of continuous renal replacement therapy and the prognosis of patients with sepsis-associated acute kidney injury]. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue. 2020;32(11):1352–1355. Chinese. 10.3760/cma.j.cn121430-20200304-00206. PMID: 33463496.

10. Lee HJ, Son YJ. Factors Associated with In-Hospital Mortality after Continuous Renal Replacement Therapy for Critically Ill Patients: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health*. 2020;17(23):8781. 10.3390/ijerph17238781. . PMID: 33256008; PMCID: PMC7730748.
11. Bagshaw SM, Darmon M, Ostermann M, Finkelstein FO, Wald R, Tolwani AJ, Goldstein SL, Gattas DJ, Uchino S, Hoste EA, Gaudry S. Current state of the art for renal replacement therapy in critically ill patients with acute kidney injury. *Intensive Care Med*. 2017;43(6):841–854. 10.1007/s00134-017-4762-8IF: 22.0 Q1. Epub 2017 Mar 13. PMID: 28289816.
12. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney Int*. 2024;105(4S):S117-S314. 10.1016/j.kint.2023.10.018. PMID: 38490803.
13. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol*. 2022;22(1):316. 10.1186/s12874-022-01801-8. . PMID: 36510134; PMCID: PMC9742671.
14. Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. *J Am Med Inf Assoc*. 2020;27(4):621–33. 10.1093/jamia/ocz228. . PMID: 32106284; PMCID: PMC7075534.
15. Bekkar M, Djemaa HK, Alitouche TA. Evaluation measures for models assessment over imbalance data sets. *J Inf Eng Appl*. 2013;3:27–38.
16. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag*. 2009;45:427–37.
17. Tsujimoto H, Tsujimoto Y, Nakata Y, Fujii T, Takahashi S, Akazawa M, Kataoka Y. Pharmacological interventions for preventing clotting of extracorporeal circuits during continuous renal replacement therapy. *Cochrane Database Syst Rev*. 2020;3:CD012467.
18. Tsujimoto Y, Miki S, Shimada H, Tsujimoto H, Yasuda H, Kataoka Y, Fujii T. Nonpharmacological interventions for preventing clotting of extracorporeal circuits during continuous renal replacement therapy. *Cochrane Database Syst Rev*. 2021;9:CD013330.
19. Zarbock A, Kullmar M, Kindgen-Milles D, et al. Effect of regional citrate anticoagulation vs systemic heparin anticoagulation during continuous kidney replacement therapy on dialysis filter life span and mortality among critically ill patients with acute kidney injury: a randomized clinical trial. *JAMA*. 2020;324:1629–39.
20. Tsujimoto Y, Fujii T. *Crit Care*. 2022;26(1):62. 10.1186/s13054-022-03910-8. . PMID: 35337352; PMCID: PMC8957143. How to Prolong Filter Life During Continuous Renal Replacement Therapy?.
21. DUNN W J SRIRAMS. Filter lifespan in critically ill adults receiving continuous renal replacement therapy: the effect of patient and treatment-related variables[J]. *Crit Care Resusc*. 2014;16(3):225–31.
22. Roberts PS. The effects of Na + and K + on the citrate activation of prothrombin and on the esterase activities of thrombokinase. *Biochim Biophys Acta*. 1970;201(2):340-9. 10.1016/0304-4165(70)90309-0. PMID: 5461512.
23. KDIGO AKI Work Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl*. 2012;2:1–138.

Figures

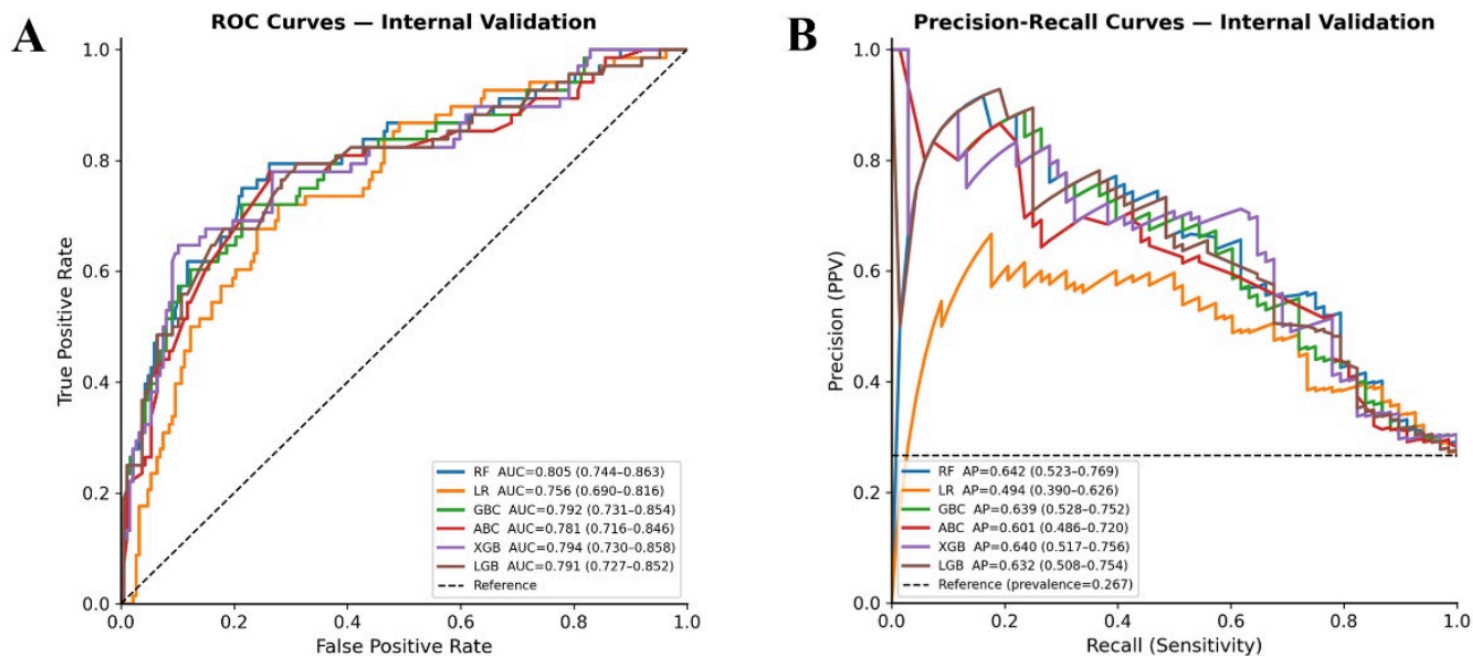


Figure 1

ROC curves (A) and PR (B) curves of the six models on the validation set

Confusion Matrices — Internal Validation

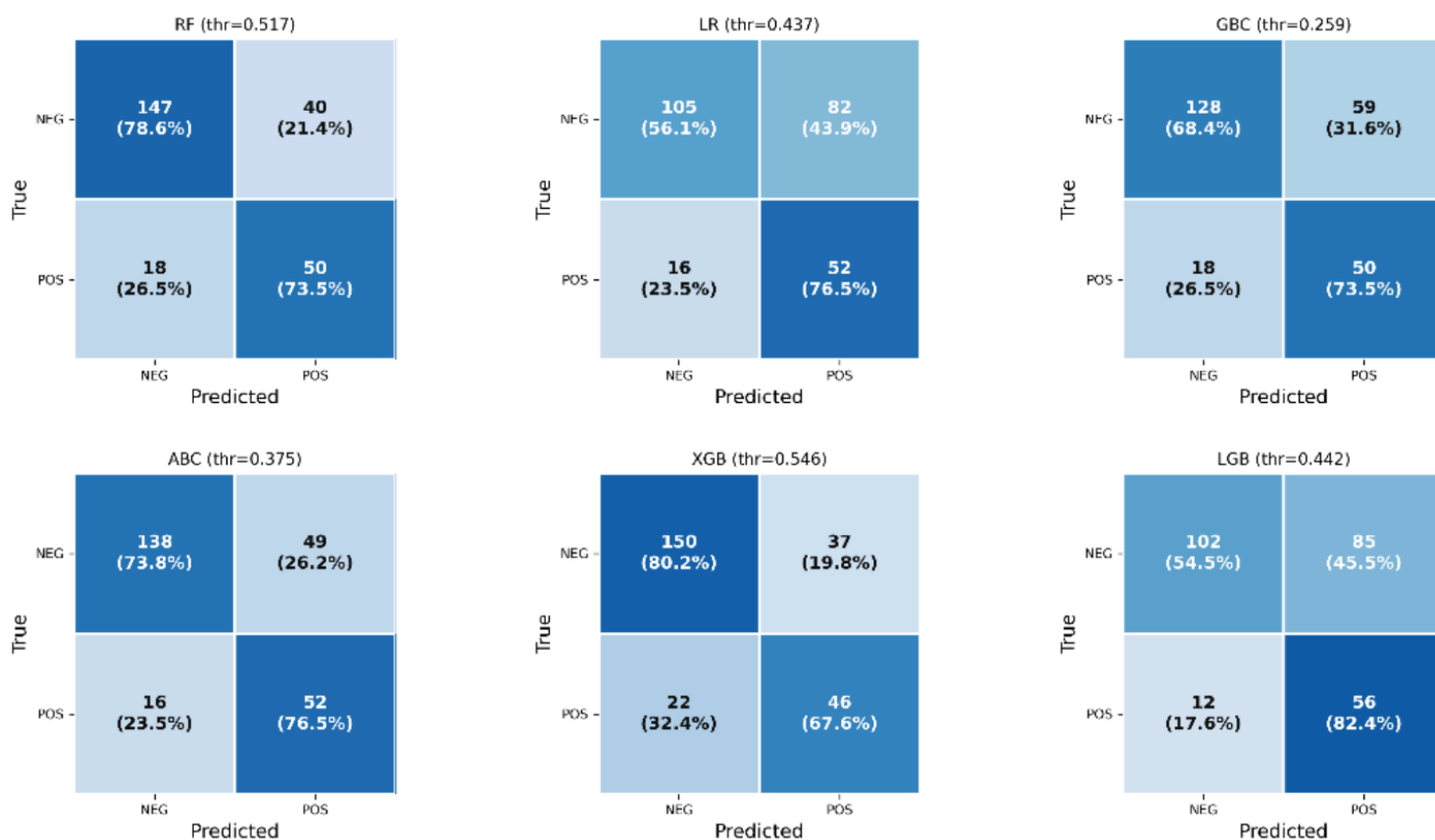


Figure 2

Confusion matrix of the random forest (RF) model on the internal validation set

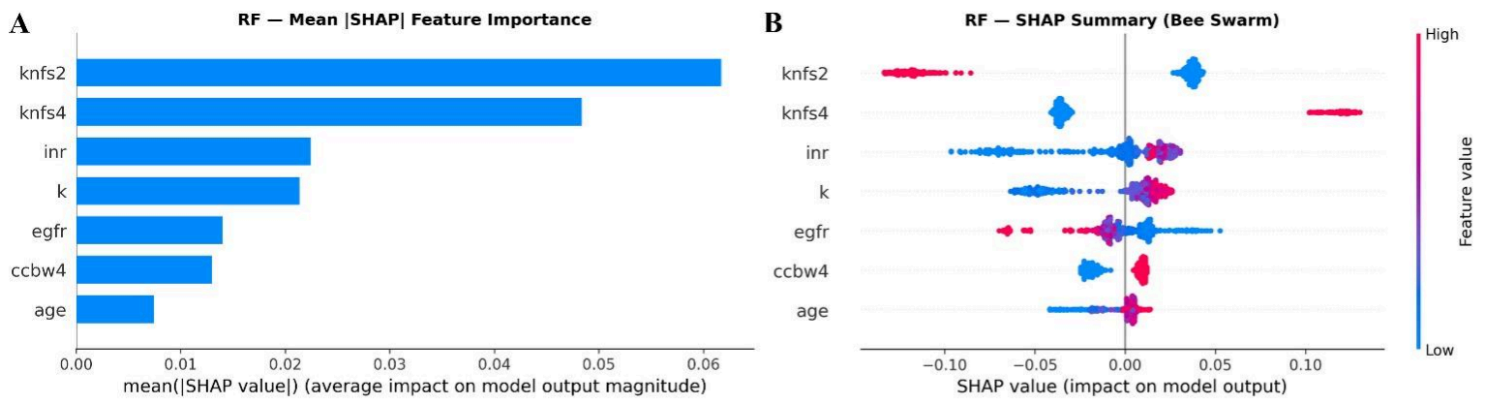


Figure 3

SHAP bar plot (A) and bee swarm (B) of feature importance for the RF model

Abbreviations: knfs2, heparin-free; knsf4, citrate; inr, INR (international normalized ratio); egfr, eGFR (estimated glomerular filtration rate); ccbw4, right internal jugular vein.

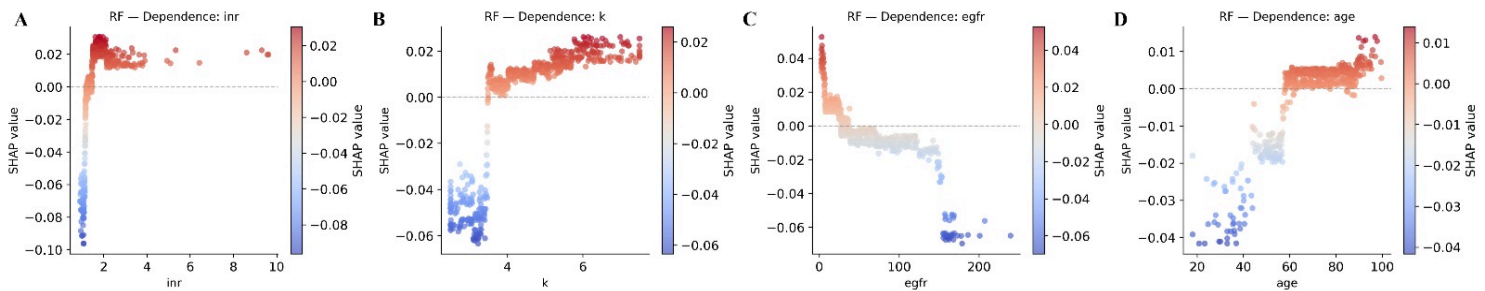


Figure 4

SHAP dependence plots for selected features in the random forest model on the internal validation set

Abbreviations: inr, INR (international normalized ratio); k, serum potassium; egfr, eGFR (estimated glomerular filtration rate);

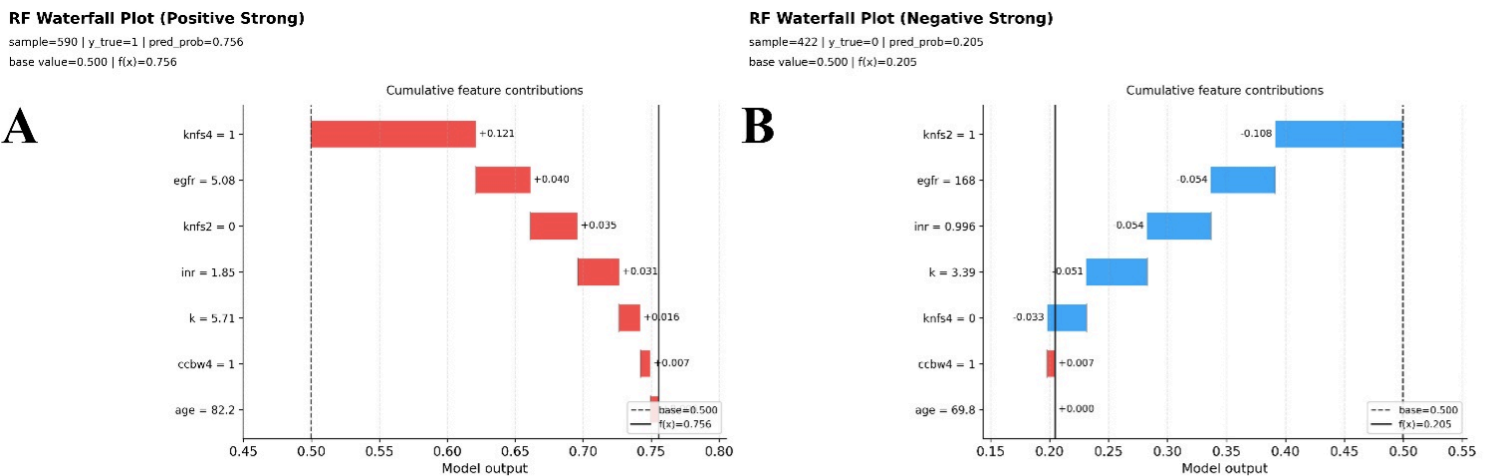


Figure 5

Individual SHAP force plot for patient A (A) and patient B (B)