Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

An Explainable Machine Learning Model to Predict Acute Kidney Injury After Cardiac Surgery: A Retrospective Cohort Study

Yuchen Gao[1,*], Chunrong Wang[2,*], Wenhao Dong[3], Bianfang Li[3], Jianhui Wang[1], Jun Li[1], Yu Tian[1], Jia Liu[1], Yuefu Wang[3]

[1]Department of Anesthesiology, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People's Republic of China; [2]Department of Anesthesiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People's Republic of China; [3]Department of Surgical Intensive Care Unit & Anesthesiology, Beijing Shijitan Hospital, Capital Medical University, Beijing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yuefu Wang, Department of Surgical Intensive Care Unit & Anesthesiology, Beijing Shijitan Hospital, Capital Medical University, 10 Tieyi Road, Haidian District, Beijing, People's Republic of China, Email wangyuefu3806@bjsjth.cn

Background: To derive and validate a machine learning (ML) prediction model of acute kidney injury (AKI) that could be used for AKI surveillance and management to improve clinical outcomes.

Methods: This retrospective cohort study was conducted in Fuwai Hospital, including patients aged 18 years and above undergoing cardiac surgery admitted between January 1, 2017, and December 31, 2018. Seventy percent of the observations were randomly selected for training and the remaining 30% for testing. The demographics, comorbidities, laboratory examination parameters, and operation details were used to construct a prediction model for AKI by logistic regression and eXtreme gradient boosting (Xgboost). The discrimination of each model was assessed on the test cohort by the area under the receiver operator characteristic (AUROC) curve, while calibration was performed by the calibration plot.

Results: A total of 15,880 patients were enrolled in this study, and 4845 (30.5%) had developed AKI. Xgboost model had the higher discriminative ability compared with logistic regression (AUROC, 0.849 [95% CI, 0.837–0.861] vs 0.803[95% CI 0.790–0.817], $P<0.001$) in the test dataset. The estimated glomerular filtration (eGFR) and creatine on intensive care unit (ICU) arrival are the two most important prediction parameters. A SHAP summary plot was used to illustrate the effects of the top 15 features attributed to the Xgboost model.

Conclusion: ML models can provide clinical decision support to determine which patients should focus on perioperative preventive treatment to preemptively reduce acute kidney injury by predicting which patients are not at risk.

Keywords: machine learning, acute kidney injury, cardiac surgery, shapley additive explanations, SHAP, prediction model

Introduction

The impact of acute kidney injury (AKI) is up to 40% in patients undergoing cardiac surgery.[1] Delayed diagnosis and intervention allow AKI to progress to more severe stages and contribute to the development of chronic kidney disease after hospital discharge.[2] Both mortality and length of stay increase with the progressive severity of kidney injury. In clinical trials, no pharmacological or non-pharmacological prevention strategies have been shown to reduce the occurrence of AKI.[3,4] Therefore, an accurate postoperative AKI risk assessment is crucial for the postoperative strategy for monitoring and disposition.

In recent years, there are many studies have attempted to predict AKI following cardiac surgery.[5] Most AKI prediction efforts focus on new blood- or urine-based biomarkers of injury, stress, and metabolomics.[6–8] In addition to being costly, these markers are not widely available and often lack sensitivity or specificity. Most prediction models

Received: 25 January 2023
Accepted: 27 September 2023
Published: 4 December 2023

currently used logistic regression, which has a general predictive ability.[9–11] With the current growth of electronic medical records coupled with machine learning presents an opportunity to improve the performance of established risk models. Compared to traditional risk scores, machine learning (ML) algorithms can easily identify nonlinear relationships and interactions between variables. ML is increasingly being used in AKI prediction, and many models use only small datasets, limiting the predictive performance of the models.[12–14] Even though a few have used large datasets, many of them suffer from the "black box" phenomenon, which could limit the acceptance among clinicians and raise ethical concerns.[15,16] SHAP is a model agnostic representation of feature importance where the impact of each feature on a particular prediction is represented using Shapley values inspired by cooperative game theory.[17] Recently, Shapley additive values as an explainable technique in artificial intelligence have been developed that may overcome the black box problem of ML models.[18,19] This new interpretable approach has been successfully applied to explain ML models related to ICU mortality prediction,[17] poststroke atrial fibrillation,[20] and postoperative complications.[21]

In this study, we aimed to demonstrate an EHR data-driven, ML approach for the prediction of postsurgical AKI. Shapley values were also reported to identify the variables that contribute most to the ML model. We hypothesize that an ML model for AKI risk prediction would outperform to a traditional logistic risk calculator.

Materials and Methods

Study Population

Consecutive adult patients who underwent cardiac surgery, admitted between January 2017 and December 2018, were recruited from Fuwai Hospital. Patients were excluded if they underwent dialysis or mechanical circulatory support before surgery. In addition, patients with missing preoperative serum creatinine level or high-level creatine (>4ng/mL) were also excluded from the analysis. The Fuwai hospital Institutional Review Board approved this project and waived consent for patients who provided research authorization compliant. We reported our work following the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement guidelines. The study was complied with the Declaration of Helsinki.

Data Collection

Baseline demographics, clinical comorbidities, and laboratory tests were extracted from the EHR. Laboratory data, including creatinine, glucose and routine blood test measured within 7 days closest to surgery were considered to be baseline values. Estimated glomerular filtration rate (eGFR) was calculated according to the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation based on serum creatinine levels.[22] Intraoperative data included transfusion, CPB data and medication were also routinely collected.

Primary Outcome

Our primary outcome of interest was AKI after surgery. AKI is diagnosed using Kidney Disease Improving Global Outcomes (KDIGO) clinical practice guidelines, based on serum creatinine: an increase in serum creatinine of 26.5 μmol/L within 48 hours after surgery or an increase at least 1.5 times the baseline measurement within 7 days after surgery. In this study, the baseline Scr of patients was measured at hospital admission.

Statistical Analyses

The normal distribution of continuous variables was determined by the Kolmogorov–Smirnov test. The continuous data were described by the median and interquartile range (IQR) or mean and standard deviation, while the categorical data were expressed by frequency and percentage. Details of the missing variables are shown in Table S1. Multiple imputations by chained equation were used to impute missing values.

Dataset was randomly divided into training set and testing set with a ratio of 7:3. We used the multivariable logistic regression (LR) with Stepwise Akaike information criterion (stepAIC) algorithm to determine the most significant variables in LR model. The ML methods eXtreme Gradient Boosting (Xgboost) and LR were used to develop and validate the models for assessing risk of AKI. The statistical details of Xgboost methods were displayed in

Supplementary methods section. Hyperparameter tuning of the Xgboost model was performed using 5-fold cross-validation during training with 70% of the data. The remaining 30% of the test set was used for validation and to compute performance results. Table S2 showed the best hyperparameters results of the Xgboost. Our final candidate model was selected based on the AUC. DeLong's test was performed to assess the differences in AUC between these two models. The calibration curve was used to compare the prediction probability of the models and the ground truth. To facilitate the interpretation of the ML model with highest AUROC, we report Shapley values. We used the SHAP values to visualize the significant features that influence the risk of AKI, to analyze the importance of individual features affecting the output of the model and to visualize the impact of key features on the final model in individuals. The caret, xgboost, SHAPforxgboost and plyr packages in R were used for XGBoost model.

R version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria) was used for statistical analysis with a 2-tailed *P* value of less than 0.05 indicating significance.

Results

Study Participants

For model development, we obtained data on 15,880 patients, of which we allocated approximately 30% of patients to the validation dataset. The flowchart of study population was presented in Figure 1. The AKI prevalence of the 15,880 patients involved in our study was 30.5%, 2.8% were stage 2 AKI, 1.6% were stage 3 AKI.

In the development dataset, the median age was 58 (IQR, 50–65) and 66.4% were male. Patients had a median body mass index of 24.9 (IQR, 22.8–27.3). The top 3 comorbidities of this cohort were hypertension (51.9%), hyperlipidemia (40.2%) and diabetes mellitus (36.0%). Coronary artery bypass grafting and valve surgery were the two most common procedures in this study. Baseline and surgical characteristics were almost well balanced between the derivation and validation datasets (Table 1). The details of derivation cohort are shown in Table 2.

Model Performance

We used the variables selected by StepAIC as input factors to develop LR and Xgboost models to predict AKI after cardiac surgery. In the training cohort, the Xgboost model exhibited a maximum AUC of 0.850 (95% CI, 0.842–0.858),

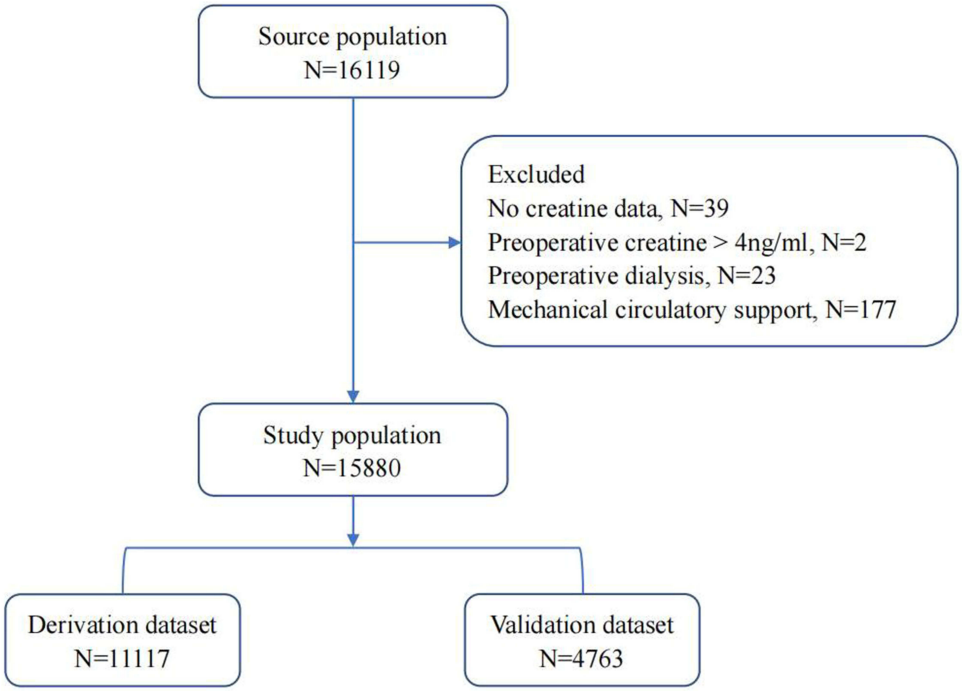

Figure 1 Study cohort description.

Table 1 Baseline Characteristics of Patients Undergoing Cardiac Surgery in Derivation and Validation Cohorts

Baseline Characteristics	Derivation Cohort N = 11,117	Validation Cohort N = 4763	*p* value	Non-AKI N=11035	AKI N=4845	*p* value
Demographic						
Age	58 [50–65]	58 [50–65]	0.623	57 [49–65]	61 [53–67]	<0.001
Sex, male	7381 (66.4)	3146 (66.1)	0.689	7306 (66.2)	3221 (66.5)	0.751
BMI	24.9 [22.8–27.3]	25.0 [22.8–27.3]	0.326	24.8 [22.8–27.1]	25.1 [22.8–27.5]	<0.001
NYHA, III/IV	4026 (36.2)	1662 (34.9)	0.116	3814 (34.6)	1874 (38.7)	<0.001
Comorbid condition						
Diabetes	3984 (35.8)	1726 (36.2)	0.643	3944 (35.7)	1766 (36.5)	0.401
Smoker	4735 (42.6)	1995 (41.9)	0.419	4649 (42.1)	2081 (43.0)	0.343
Peripheral vascular disease	873 (7.9)	370 (7.8)	0.881	789 (7.2)	454 (9.4)	<0.001
Heart failure	2967 (26.7)	1232 (25.9)	0.290	2620 (23.7)	1579 (32.6)	<0.001
Hypertension	5731 (51.6)	2504 (52.6)	0.245	5323 (48.2)	2912 (60.1)	<0.001
Cerebrovascular disease	2958 (26.6)	1229 (25.8)	0.301	2803 (25.4)	1384 (28.6)	<0.001
COPD	165 (1.5)	71 (1.5)	1	142 (1.3)	94 (1.9)	0.002
Previous heart surgery	489 (4.4)	200 (4.2)	0.601	382 (3.5)	307 (6.3)	<0.001
Myocardial infarction	1844 (16.6)	778 (16.3)	0.711	1857 (16.8)	765 (15.8)	0.110
Hyperlipidemia	4467 (40.2)	1920 (40.3)	0.893	4490 (40.7)	1897 (39.2)	0.072
Anemia	2520 (22.7)	1102 (23.1)	0.532	2227 (20.2)	1395 (28.8)	<0.001
Chronic kidney disease	1185 (10.7)	519 (10.9)	0.679	891 (8.1)	813 (16.8)	<0.001
Atrial fibrillation	1531 (13.8)	613 (12.9)	0.134	1252 (11.3)	892 (18.4)	<0.001
Infective endocarditis	106 (1.0)	41 (0.9)	0.639	102 (0.9)	45 (0.9)	1
Pulmonary hypertension	1497 (13.5)	656 (13.8)	0.622	1389 (12.6)	764 (15.8)	<0.001
Preoperative laboratory variables						
Glu, mmol/L	5.1 [4.6–6.1]	5.2 [4.6–6.2]	0.126	5.1 [4.6–6.1]	5.2 [4.6–6.3]	<0.001
NT-proBNP, pg/mL	247.0 [83.4–805.7]	241.0 [82.0–791.3]	0.374	194.4 [70.8–633.6]	416.3 [131.9–1155]	<0.001
K, mmol/L	4.0 [3.8–4.3]	4.0 [3.8–4.3]	0.625	4.02 [3.81–4.27]	4.02 [3.78–4.29]	0.828
Cl, mmol/L	104.7 [102.4–106.6]	104.6 [102.3–106.6]	0.393	104.6 [102.3–106.5]	104.8 [102.3–106.7]	0.031
Na, mmol/L	140.2 [138.2–142.2]	140.2 [138.1–142.2]	0.434	140.2 [138.2–142.2]	140.2 [138.1–142.2]	0.434
Ca, mmol/L	2.3 [2.2–2.3]	2.3 [2.2–2.3]	0385	2.25 [2.18–2.33]	2.25 [2.18–2.32]	0.385
AST, IU/L	24 [20–30]	24 [19–30]	0.531	24 [20–30]	24 [20–31]	0.358
ALP, IU/L	64 [53–78]	64 [53–78]	0.895	64 [53–78]	65 [54–79]	<0.001
TBil, μmol/L	11.5 [8.6–15.8]	11.4 [8.5–15.6]	0.092	11.4 [8.6–15.4]	11.7 [8.5–16.5]	0.001
WBC, 10^9/L	6.24 [5.22–7.47]	6.29 [5.23–7.53]	0.296	6.24 [5.22–7.44]	6.28 [5.23–7.58]	0.016
PLT, 10^9/L	201 [167–242]	203 [167–242]	0.538	205 [171–244]	196 [159–237]	<0.001
Creatinine, μmol/L	82.9 [72.0–94.0]	82.0 [72.0–94.0]	0.494	82.0 [72.0–93.0]	84.0 [72.8–97.6]	<0.001
eGFR, mL/min/1.73m^2	83.3 [70.9–97.3]	83.7 [70.7–97.0]	0.676	91.4 [77.2–110.5]	89.3 [73.2–106.5]	<0.001

Albumin, g/L	40.1 [37.8–42.6]	40.2 [37.9–42.8]	0.059	40.4 [38.1–42.9]	39.6 [37.2–42.1]	<0.001
HsCRP, mg/L	1.15 [0.50–2.79]	1.19 [0.50–2.9]	0.306	1.04 [0.45–2.48]	1.51 [0.64–3.70]	<0.001
HDL-C, mmol/L	1.09 [0.92–1.31]	1.09 [0.91–1.31]	0.822	1.10 [0.92–1.32]	1.07 [0.89–1.29]	<0.001
LDL-C, mmol/L	2.37 [1.87–3.00]	2.37 [1.87–2.98]	0.400	2.35 [1.85–2.99]	2.38 [1.90–3.00]	0.022
TG, mmol/L	1.28 [0.95–1.80]	1.29 [0.95–1.80]	0.599	1.28 [0.94–1.78]	1.30 [0.96–1.86]	0.009
Apolipoprotein A1, g/L	1.24 [1.09–1.41]	1.23 [1.09–1.40]	0.656	1.24 [1.09–1.41]	1.22 [1.08–1.40]	<0.001
Preoperative Echo						
LVEF, %	62 [59–65]	62 [59–65]	0.855	62 [60–65]	61 [58–65]	<0.001
LVRWT	0.38 [0.33–0.43]	0.38 [0.33–0.43]	0.736	0.38 [0.33–0.43]	0.38 [0.33–0.43]	0.001
Surgical characteristics						
Emergent	180 (2.3)	240 (7.1)	<0.001	255 (2.3)	325 (6.7)	<0.001
CPB or not	8675 (78.0)	3727 (78.2)	0.780	8285 (75.1)	4117 (85.0)	<0.001
CPB time, min	106 [81–140]	106 [82–138]	0.747	101 [78–131]	121 [90–159]	<0.001
Operation time, min	286[245–340]	286[245–340]	0.972	280 [240–326]	309 [259–375]	<0.001
Cross-clamp time, min	75 [55–102]	75 [54–101]	0.731	71 [52–96]	83 [60–112]	<0.001
Type of surgery			0.548			<0.001
CABG surgery	5084 (45.7)	2182 (45.8)		5288 (47.9)	1978 (40.8)	
Valve surgery	3348 (30.1)	1453 (30.5)		3130 (28.4)	1671 (34.5)	
Aortic surgery	1264 (11.4)	517 (10.9)		1027 (9.3)	754 (15.6)	
CHD surgery	874 (7.9)	394 (8.3)		1010 (9.2)	258 (5.3)	
Cardiomyopathy surgery	419 (3.8)	175 (3.7)		439 (4.0)	155 (3.2)	
Cardiac tumor surgery	128 (1.2)	42 (0.9)		141 (1.3)	29 (0.6)	
Intraoperative characteristics						
RBC use	5240 (47.1)	2218 (46.6)	0.522	4798 (43.5)	2660 (54.9)	<0.001
Plasma use	1435 (12.9)	601 (12.6)	0.635	1024 (9.3)	1012 (20.9)	<0.001
Corticosteroids	205 (1.8)	87 (1.8)	0.992	122 (1.1)	170 (3.5)	<0.001
Levosimendan	2669 (24.0)	1136 (23.9)	0.847	2438 (22.1)	1367 (28.2)	<0.001
Colloidal solution	4149 (37.3)	1740 (36.5)	0.354	4132 (37.4)	1757 (36.3)	0.162
Postoperative laboratory variables						
Creatinine ICU arrival	77.2 [66.4–90.7]	77.1 [66.4–90.6]	0.995	74.0 [64.2–85.1]	87.4 [73.9–104.7]	<0.001
eGFR ICU arrival	89.6 [74.3–106.9]	89.8 [74.4–106.8]	0.924	95.2 [81.1–111.6]	75.9 [61.7–91.6]	<0.001
NT-proBNP	741.4 [420.4–1347.0]	716.5 [403.4–1351.0]	0.128	664.9 [384.9–1167]	945.2 [517.9–1855]	<0.001

Abbreviations: AKI, acute kidney injury; BMI, body mass index; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease; Glu, glucose; NT-proBNP, N-terminal pro-B-type natriuretic peptide; K, potassium; Cl, chlorine; Na, sodium; Ca, calcium; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBil, total bilirubin; WBC, white blood cell; PLT, platelet; eGFR, estimated glomerular filtration rate; HsCRP, hypersensitive C reactive protein; HDL-C, high-density lipoprotein-cholesterol; LDL-C, low-density lipoprotein-cholesterol; TG, triglyceride; LVEF, left ventricular ejection fraction; LVRWT, left ventricular relative wall thickness; CPB, cardiopulmonary bypass; CABG, coronary artery bypass graft; CHD, congenital heart disease; RBC, red blood cell; ICU, intensive care unit.

https://doi.org/10.2147/CLEP.S404580

Table 2 Baseline Characteristics in Derivation Cohort

Baseline Characteristics	Non-AKI N=7725	AKI N=3392	*p* value
Demographic			
Age	57 [49–65]	61 [53–67]	<0.001
Sex, male	5130(66.4)	2251 (66.4)	0.980
BMI	24.8 [22.8–27.1]	25.1 [22.8–27.5]	0.015
NYHA, III/IV	2716 (35.2)	1310 (38.6)	0.001
Comorbid condition			
Diabetes	2762 (35.8)	1222 (36.0)	0.800
Smoker	3272 (42.4)	1463 (43.1)	0.459
Peripheral vascular disease	548 (7.1)	325 (9.6)	<0.001
Heart failure	1856 (24.0)	1111 (32.8)	<0.001
Hypertension	3724 (48.2)	2007 (59.2)	<0.001
Cerebrovascular disease	1975 (25.6)	983 (29.0)	<0.001
COPD	104 (1.3)	61 (1.8)	0.084
Previous heart surgery	265 (3.4)	224 (6.6)	<0.001
Myocardial infarction	1307 (16.9)	537 (15.8)	0.164
Hyperlipidemia	3139 (40.6)	1328 (39.2)	0.148
Anemia	1531 (19.8)	989 (29.2)	<0.001
Chronic kidney disease	605 (7.8)	580 (17.1)	<0.001
Atrial fibrillation	896 (11.6)	635 (18.7)	<0.001
Infective endocarditis	74 (1.0)	32 (0.9)	1
Pulmonary hypertension	971 (12.6)	526 (15.5)	<0.001
Preoperative laboratory variables			
Glu, mmol/L	5.1 [4.6–6.1]	5.2 [4.6–6.3]	0.001
NT-proBNP, pg/mL	194.8 [71.7–625.8]	428.0 [133.1–1191.3]	<0.001
K, mmol/L	4.03 [3.81–4.27]	4.02 [3.78–4.29]	0.484
Cl, mmol/L	104.6 [102.4–106.5]	104.8 [102.3–106.7]	0.066
Na, mmol/L	140.4 [138.4–142.3]	140.0 [138.0–141.9]	<0.001
Ca, mmol/L	2.25 [2.18–2.33]	2.24 [2.17–2.32]	<0.001
AST, IU/L	24 [20–30]	24 [20–31]	0.849
ALP, IU/L	64 [53–77]	65 [54–79]	<0.001
TBil, μmol/L	11.4 [8.6–15.4]	11.8 [8.6–16.9]	<0.001
WBC, 10^9/L	6.24 [5.22–7.44]	6.27 [5.21–7.55]	0.113
PLT, 10^9/L	205 [171–244]	193 [159–235]	<0.001
Creatinine, μmol/L	74.0 [64.1–85.0]	87.6 [74.0–104.4]	<0.001
eGFR, mL/min/1.73m^2	84.7 [72.8–98.3]	80.2 [65.8–94.7]	<0.001
Albumin, g/L	40.3 [38.1–42.8]	39.5 [37.2–42.1]	<0.001
HsCRP, mg/L	1.04 [0.45–2.44]	1.53 [0.63–3.65]	<0.001
HDL-C, mmol/L	1.10 [0.92–1.32]	1.07 [0.90–1.28]	<0.001
LDL-C, mmol/L	2.36 [1.85–2.99]	2.40 [1.90–3.00]	0.03
TG, mmol/L	1.28 [0.94–1.78]	1.29 [0.95–1.86]	0.099
Apolipoprotein A1, g/L	1.24 [1.09–1.41]	1.22 [1.08–1.40]	0.002
Preoperative Echo			
LVEF, %	62 [60–65]	60 [58–65]	<0.001
LVRWT	0.38 [0.33–0.43]	0.38 [0.33–0.43]	0.026
Surgical characteristics			
Emergent	180 (2.3)	240 (7.1)	<0.001
CPB or not	5792 (75.0)	2883 (85.0)	<0.001
CPB time, min	100 [78–131]	121 [90–160]	<0.001
Operation time, min	280 [240–325]	310 [260–376]	<0.001
Cross-clamp time, min	71 [52–96]	83 [60–112]	<0.001

(*Continued*)

Table 2 (Continued).

Baseline Characteristics	Non-AKI N=7725	AKI N=3392	*p* value
Type of surgery			<0.001
CABG surgery	3720 (48.2)	1364 (40.2)	
Valve surgery	2174 (28.1)	1174 (34.6)	
Aortic surgery	720 (9.3)	544 (16.0)	
CHD surgery	696 (9.0)	178 (5.2)	
Cardiomyopathy surgery	309 (4.0)	110 (3.2)	
Cardiac tumor surgery	106 (1.4)	22 (0.6)	
Intraoperative characteristics			
RBC use	3374 (43.7)	1866 (55.0)	<0.001
Plasma use	714 (9.2)	721 (21.3)	<0.001
Corticosteroids	84 (1.1)	121 (3.6)	<0.001
Levosimendan	1713 (22.2)	956 (28.2)	<0.001
Colloidal solution	2909 (37.7)	1240 (36.6)	0.279
Postoperative laboratory variables			
Creatinine ICU arrival	74.0 [64.1–85.0]	87.6 [74.0–104.4]	<0.001
eGFR ICU arrival	95.2 [81.1–111.8]	75.9 [61.9–91.5]	<0.001
NT-proBNP	670.3 [390.7–1156.0]	965.9 [529.9–1876.0]	<0.001

Abbreviations: AKI, acute kidney injury; BMI, body mass index; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease; Glu, glucose; NT-proBNP, N-terminal pro-B-type natriuretic peptide; K, potassium; Cl, chlorine; Na, sodium; Ca, calcium; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBil, total bilirubin; WBC, white blood cell; PLT, platelet; eGFR, estimated glomerular filtration rate; HsCRP, hypersensitive C reactive protein; HDL-C, high-density lipoprotein-cholesterol; LDL-C, low-density lipoprotein-cholesterol; TG, triglyceride; LVEF, left ventricular ejection fraction; LVRWT, left ventricular relative wall thickness; CPB, cardiopulmonary bypass; CABG, coronary artery bypass graft; CHD, congenital heart disease; RBC, red blood cell; ICU, intensive care unit.

which performed significantly better than LR of 0.807 (95% CI, 0.798–0.816, $p<0.001$). Overall, the predictive ability of models was great in the validation dataset with AUC of was 0.849 (95% CI, 0.837–0.861) in Xgboost, and 0.803 (95% CI, 0.790–0.817) in LR (Table 3). The performance of these two models on the derivation and validation cohorts is proved in Figure 2. Calibration curves were plotted to show the relationship between the predictions of these two models and the observations of the cohort, where a fully calibrated model follows a 45° line (Figure 3). The variance inflation factor and variable contribution of the LR model are displayed in Table S3 and Figure S1 respectively.

Model Explainability

The contribution of the top 15 feature in the model was presented by the SHAP summary plot (Figure 4a). The plots also identify the features that influenced the model predictions the most. Figure 4b shows the variable importance in the Xgboost model, which was the best performing model for postoperative AKI. Additionally, SHAP dependence analysis was performed to depict how a single predictor affected the output of the Xgboost model. The top six clinical features that have the most important impact on the output of the Xgboost prediction model have been shown in detail in Figure 5.

Table 3 Prediction Performance of the Models

Model	Training Set		Test Set	
	AUROC	*p* value*	AUROC	*p* value*
Logistic regression	0.807(0.798–0.816)	<0.001	0.803(0.790–0.817)	<0.001
Xgboost	0.850(0.842–0.858)		0.849(0.837–0.861)	

Note: *Indicates the difference in C-index compared with the logistic regression model.
Abbreviations: Xgboost, eXtreme gradient boosting; AUROC, area under the receiver operator characteristic curve.

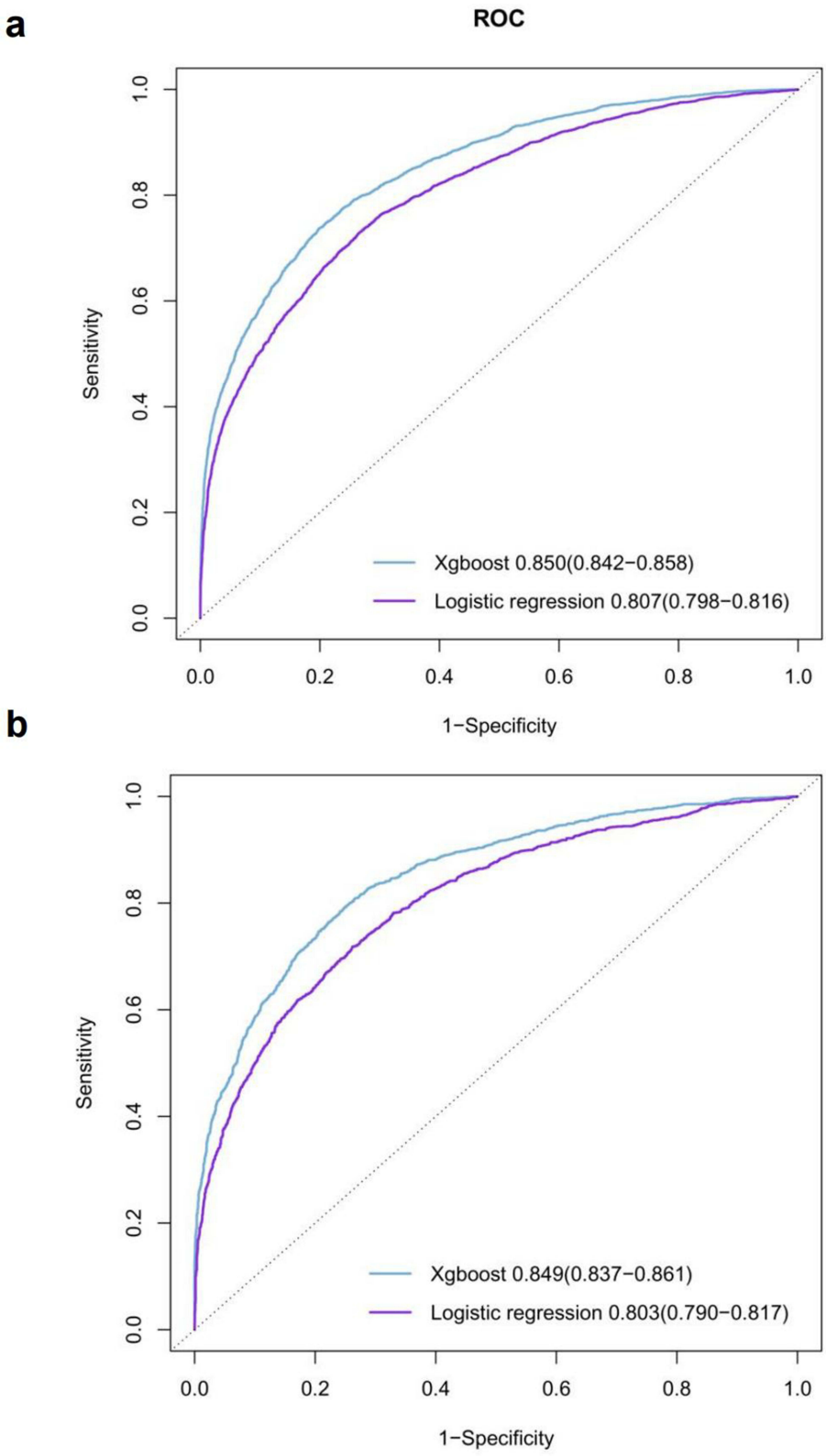

Figure 2 Comparison of AUC between LR and Xgboost models in predicting postoperative AKI. (**a**) derivation cohort. (**b**) testing cohort.
Abbreviations: AUC, area under curve; Xgboost, eXtreme gradient boosting; LR, logistic regression.

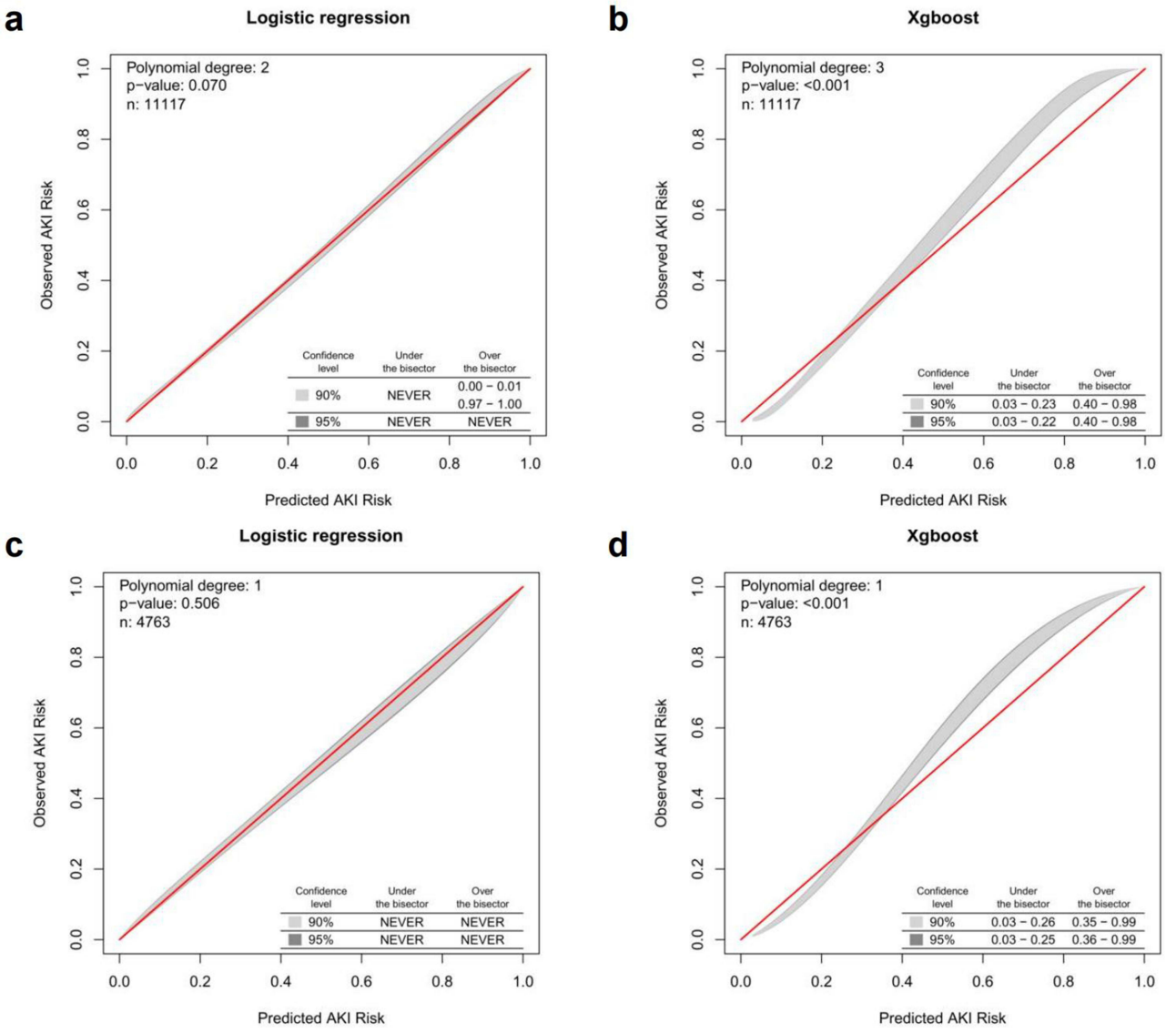

Figure 3 Calibration plot of Logistic regression ((**a**) derivation cohort, (**c**) testing cohort) and Xgboost model ((**b**) derivation cohort, (**d**) testing cohort).

Discussion

Data-driven prediction is not new in medicine and is already routinely used in clinical practice. We developed and validated a machine learning model with postoperative AKI as the outcome and evaluated the performance in this cohort study. ML model demonstrated excellent performance in predicting AKI among individuals with cardiac surgery. The most significant advancement with our risk models is the use of explainable ML approach to risk prediction. This study contributes to the growing body of retrospective machine learning literature for the prediction of AKI.

Prior risk prediction models have been limited to revealing the linear relationship between features and AKI.[11,23] Besides traditional AKI prediction models, the number of studies applying ML to predict AKI has grown steadily over the past decade.[24] However, these studies have been mainly developed in small cohorts, and absence of model interpretability.[12,16] To this point, the current study was important because it proposed and internally validated a ML algorithm to predict the risk of AKI based on a large prospectively collected dataset focused on a population of patients undergoing cardiac surgery.

The performance of ML algorithms against conventional statistical methods such as LR to predict clinical outcome remains controversial, some studies showed that both models performed equally well,[25,26] while other studies reported better performance with ML models.[27,28] Luo et al[29] were the first to apply machine learning approaches to predict the

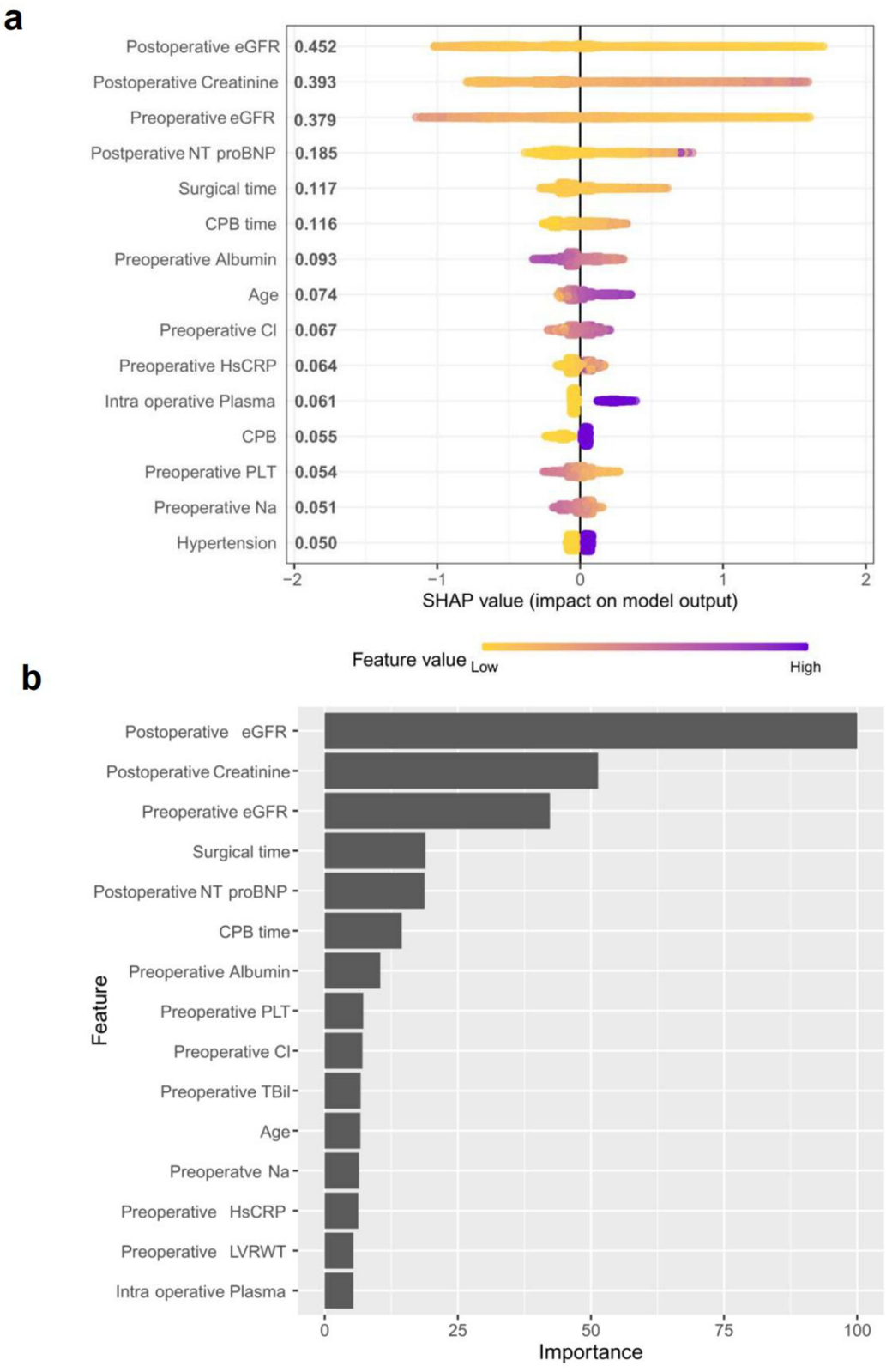

Figure 4 (**a**) sHapley Additive exPlanations (SHAP) summary plot of the top 15 features in the Xgboost model. The higher the SHAP value of a feature, the higher the probability of postoperative AKI development. Each line represents a feature, and a single dot represent each value for each variable observed in the cohort. Purple represents higher feature values, and yellow represents lower feature values. (**b**) Variable importance of features included in the Xgboost model for prediction of AKI.
Abbreviations: AKI, acute kidney injury; eGFR, estimated glomerular filtration rate; NT-proBNP, N-terminal pro-B-type natriuretic peptide; CPB, cardiopulmonary bypass; LVRWT, left ventricular relative wall thickness.

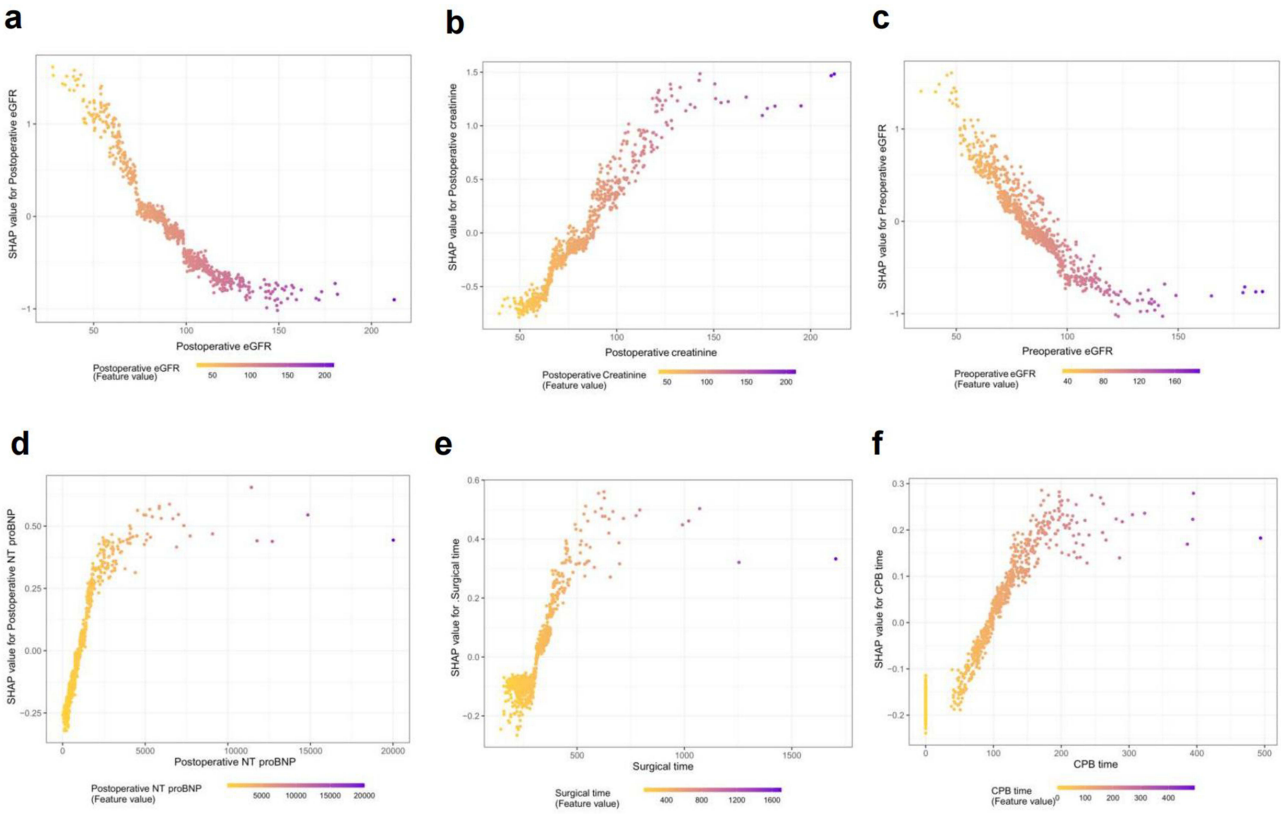

Figure 5 SHAP dependence plot of the Xgboost model. The SHAP dependence plot shows how a single feature affects the output of the Xgboost prediction model. SHAP values for specific features exceed zero, representing an increased risk of AKI development. (**a**) Postoperative eGFR; (**b**) Postoperative creatinine; (**c**) Preoperative eGFR; (**d**) Postoperative NT-proBNP; (**e**) Surgical time; (**f**) CPB time.
Abbreviations: eGFR, estimated glomerular filtration rate; SHAP, shapley additive explanations.

cardiac surgery-associated acute kidney injury (CSA-AKI) in pediatric patients undergoing cardiac surgery. The Xgboost model achieved the best predictive performance among the considered machine learning models. The results of external validation and sensitivity analysis demonstrated the robustness and applicability of the Xgboost model for predicting pediatric postoperative cardiac AKI. Our study further confirmed the excellent performance of the Xgboost model in predicting postoperative AKI in adults undergoing cardiac surgery.

Many of the most predictive features (eg, creatinine, eGFR and surgical time) identified in previous work were confirmed in the models presented here. Except preoperative and intraoperative characteristics, we also included a number of post-operative laboratory tests at ICU arrival, which may provide some additional information on the kidney function of patients in the early postoperative period. Additionally, our ML model reflects complex nonlinear relationships, which can assist in comprehension the association between the changes in predictors and the risk of AKI. A growing amount of research suggests that baseline NT-proBNP can accurately predict AKI in both surgical and medical patient.[30–32] We found that postoperative NT-proBNP can also have a contribution of AKI development. Almost 14% patients without preoperative NT-proBNP level in our study, we can speculate that early postoperative NT-proBNP assessment may contribute to the risk prediction of AKI. Early postoperative prediction helps to optimize postoperative management and care planning, such as continuous assessment of renal function, hemodynamic monitoring, avoidance of renal toxins, or renal replacement therapy. This study used preoperative, intraoperative, and postoperative variables routinely collected in clinical practice and does not add additional laboratory tests and financial burden to standard clinical procedures.

SHAP analysis was employed in ML application, providing a visualizable prediction model that can be easily accepted by clinicians or decision makers. The algorithm of model interpretability that allowed a quick understanding of the impact of single features on model predictions. In the present study, the SHAP method was used to interpret the XGBoost model, and the results showed that eGFR, creatinine and NT-proBNP at the time of ICU admission were the three most important variables in

predicting postoperative cardiac AKI. Our results have emphasized the essentiality of early postoperative laboratory biomarkers in reflecting the acute pathophysiology of kidney injury. Notably, 80% of the top 5 most important predictor variables were intraoperative and postoperative variables, suggesting a significant effect of intraoperative management and surgical operations on the occurrence of AKI.

Several limitations should be mentioned in this study. First, this study is retrospective and relates to only one center, although our dataset was relatively large, complete and accurate. Study findings may not be generalizable for other settings. Second, some residual confounders such as intraoperative time-series variables are unrecorded, may have biased the results. Third, we did not use urine output to define AKI because it was not available to most patients. The lack of urine output criteria may have resulted in lower prevalence of AKI than previous studies. However, the clinical use of diuretics will affect this diagnostic criterion. Lastly, we encourage validation of the findings and algorithm of this study in other cohort, as well as the inclusion of novel peri-operative data types in the analysis. Future research will shift from risk stratification to therapeutic interventions, which will be a milestone in clinical practice

Conclusion

A robust model for predicting AKI after cardiac surgery has been obtained using machine learning techniques. This tool can provide relevant information to patients and assists in clinical decision-making process. However, the model still needs to be further verified by external validation or a randomized clinical trial before using in the daily clinical practice.

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Funding

This work is supported by the Chinese Academy of Medical Sciences Central Public Welfare Scientific Research Institute Basal Research Expenses (2019XK320052).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Fuhrman DY, Kellum JA. Epidemiology and pathophysiology of cardiac surgery-associated acute kidney injury. *Curr Opin Anaesthesiol*. 2017;30(1):60–65. doi:10.1097/ACO.0000000000000412
2. Himmelfarb J, Joannidis M, Molitoris B, et al. Evaluation and initial management of acute kidney injury. *Clin J Am Soc Nephrol*. 2008;3(4):962–967. doi:10.2215/CJN.04971107
3. Hausenloy DJ, Candilio L, Evans R, et al. Remote ischemic preconditioning and outcomes of cardiac surgery. *N Engl J Med*. 2015;373(15):1408–1417. doi:10.1056/NEJMoa1413534
4. Billings FT, Hendricks PA, Schildcrout JS, et al. High-dose perioperative atorvastatin and acute kidney injury following cardiac surgery: a randomized clinical trial. *JAMA*. 2016;315(9):877–888. doi:10.1001/jama.2016.0548
5. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. *Ann Thorac Surg*. 2012;93(1):337–347. doi:10.1016/j.athoracsur.2011.09.010
6. Wen Y, Parikh CR. Current concepts and advances in biomarkers of acute kidney injury. *Crit Rev Clin Lab Sci*. 2021;58(5):354–368. doi:10.1080/10408363.2021.1879000
7. Kashani K, Cheungpasitporn W, Ronco C. Biomarkers of acute kidney injury: the pathway from discovery to clinical adoption. *Clin Chem Lab Med*. 2017;55(8):1074–1089. doi:10.1515/cclm-2016-0973
8. Fox BM, Gil HW, Kirkbride-Romeo L, et al. Metabolomics assessment reveals oxidative stress and altered energy production in the heart after ischemic acute kidney injury in mice. *Kidney Int*. 2019;95(3):590–610. doi:10.1016/j.kint.2018.10.020
9. Guan C, Li C, Xu L, et al. Risk factors of cardiac surgery-associated acute kidney injury: development and validation of a perioperative predictive nomogram. *J Nephrol*. 2019;32(6):937–945. doi:10.1007/s40620-019-00624-z
10. Coulson T, Bailey M, Pilcher D, et al. Predicting acute kidney injury after cardiac surgery using a simpler model. *J Cardiothorac Vasc Anesth*. 2021;35(3):866–873. doi:10.1053/j.jvca.2020.06.072
11. Thakar CV, Arrigain S, Worley S, Yared JP, Paganini EP. A clinical score to predict acute renal failure after cardiac surgery. *J Am Soc Nephrol*. 2005;16(1):162–168. doi:10.1681/ASN.2004040331

12. Tseng PY, Chen YT, Wang CH, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care*. 2020;24(1):478. doi:10.1186/s13054-020-03179-9
13. Kalisnik JM, Bauer A, Vogt FA, et al. Artificial intelligence-based early detection of acute kidney injury after cardiac surgery. *Eur J Cardiothorac Surg*. 2022;62(5). doi:10.1093/ejcts/ezac289.
14. Petrosyan Y, Mesana TG, Sun LY. Prediction of acute kidney injury risk after cardiac surgery: using a hybrid machine learning algorithm. *BMC Med Inform Decis Mak*. 2022;22(1):137. doi:10.1186/s12911-022-01859-w
15. Koyner JL, Carey KA, Edelson DP, Churpek MM. The development of a machine learning inpatient acute kidney injury prediction model. *Crit Care Med*. 2018;46(7):1070–1077. doi:10.1097/CCM.0000000000003123
16. Lee HC, Yoon HK, Nam K, et al. Derivation and validation of machine learning approaches to predict acute kidney injury after cardiac surgery. *J Clin Med*. 2018;7(10):322.
17. Thorsen-Meyer HC, Nielsen AB, Nielsen AP, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Digit Health*. 2020;2(4):e179–e191. doi:10.1016/S2589-7500(20)30018-2
18. Ribeiro MT, Singh S, Guestrin C. Model-Agnostic Interpretability of Machine Learning; 2016.
19. Rodríguez-Pérez R, Bajorath J. Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. *J Med Chem*. 2020;63(16):8761–8777. doi:10.1021/acs.jmedchem.9b01101
20. Zheng X, Wang F, Zhang J, et al. Using machine learning to predict atrial fibrillation diagnosed after ischemic stroke. *Int J Cardiol*. 2022;347:21–27. doi:10.1016/j.ijcard.2021.11.005
21. Xue B, Li D, Lu C, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw Open*. 2021;4(3):e212240. doi:10.1001/jamanetworkopen.2021.2240
22. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604–612. doi:10.7326/0003-4819-150-9-200905050-00006
23. Palomba H, de Castro I, Neto AL, Lage S, Yu L. Acute kidney injury prediction following elective cardiac surgery: AKICS Score. *Kidney Int*. 2007;72(5):624–631. doi:10.1038/sj.ki.5002419
24. Vagliano I, Chesnaye NC, Leopold JH, Jager KJ, Abu-Hanna A, Schut MC. Machine learning models for predicting acute kidney injury: a systematic review and critical appraisal. *Clin Kidney J*. 2022;15(12):2266–2280. doi:10.1093/ckj/sfac181
25. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12–22. doi:10.1016/j.jclinepi.2019.02.004
26. Penny-Dimri JC, Bergmeir C, Perry L, Hayes L, Bellomo R, Smith JA. Machine learning to predict adverse outcomes after cardiac surgery: a systematic review and meta-analysis. *J Card Surg*. 2022;37(11):3838–3845. doi:10.1111/jocs.16842
27. Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intens Care Med*. 2020;46(3):383–400. doi:10.1007/s00134-019-05872-y
28. Islam MM, Nasrin T, Walther BA, Wu CC, Yang HC, Li YC. Prediction of sepsis patients using machine learning approach: a meta-analysis. *Comput Methods Programs Biomed*. 2019;170:1–9. doi:10.1016/j.cmpb.2018.12.027
29. Luo XQ, Kang YX, Duan SB, et al. Machine learning-based prediction of acute kidney injury following pediatric cardiac surgery: model development and validation study. *J Med Internet Res*. 2023;25:e41142.
30. Belley-Côté EP, Parikh CR, Shortt CR, et al. Association of cardiac biomarkers with acute kidney injury after cardiac surgery: a multicenter cohort study. *J Thorac Cardiovasc Surg*. 2016;152(1):245–251.e4. doi:10.1016/j.jtcvs.2016.02.029
31. Verwijmeren L, Bosma M, Vernooij LM, et al. Associations between preoperative biomarkers and cardiac surgery-associated acute kidney injury in elderly patients: a cohort study. *Anesth Analg*. 2021;133(3):570–577. doi:10.1213/ANE.0000000000005650
32. Naruse H, Ishii J, Takahashi H, et al. Predicting acute kidney injury using urinary liver-type fatty-acid binding protein and serum N-terminal pro-B-type natriuretic peptide levels in patients treated at medical cardiac intensive care units. *Crit Care*. 2018;22(1):197. doi:10.1186/s13054-018-2120-z