Retrospective Cohort Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Development of interpretable machine learning models for prediction of acute kidney injury after noncardiac surgery: a retrospective cohort study

Rao Sun, PhD[a], Shiyong Li, PhD[a], Yuna Wei, MD[c], Liu Hu, MD[b], Qiaoqiao Xu, PhD[a], Gaofeng Zhan, PhD[a], Xu Yan, MD[a], Yuqin He, MD[a], Yao Wang, MSc[c], Xinhua Li, MD[a,*], Ailin Luo, PhD[a,*], Zhiqiang Zhou, PhD[a,*]

Background: Early identification of patients at high-risk of postoperative acute kidney injury (AKI) can facilitate the development of preventive approaches. This study aimed to develop prediction models for postoperative AKI in noncardiac surgery using machine learning algorithms. The authors also evaluated the predictive performance of models that included only preoperative variables or only important predictors.

Materials and methods: Adult patients undergoing noncardiac surgery were retrospectively included in the study (76 457 patients in the discovery cohort and 11 910 patients in the validation cohort). AKI was determined using the KDIGO criteria. The prediction model was developed using 87 variables (56 preoperative variables and 31 intraoperative variables). A variety of machine learning algorithms were employed to develop the model, including logistic regression, random forest, extreme gradient boosting, and gradient boosting decision trees. The performance of different models was compared using the area under the receiver operating characteristic curve (AUROC). Shapley Additive Explanations (SHAP) analysis was employed for model interpretation.

Results: The patients in the discovery cohort had a median age of 52 years (IQR: 42–61 years), and 1179 patients (1.5%) developed AKI after surgery. The gradient boosting decision trees algorithm showed the best predictive performance using all available variables, or only preoperative variables. The AUROCs were 0.849 (95% CI: 0.835–0.863) and 0.828 (95% CI: 0.813–0.843), respectively. The SHAP analysis showed that age, surgical duration, preoperative serum creatinine, and gamma-glutamyltransferase, as well as American Society of Anesthesiologists physical status III were the most important five features. When gradually reducing the features, the AUROCs decreased from 0.852 (including the top 40 features) to 0.839 (including the top 10 features). In the validation cohort, the authors observed a similar pattern regarding the models' predictive performance.

Conclusions: The machine learning models the authors developed had satisfactory predictive performance for identifying high-risk postoperative AKI patients. Furthermore, the authors found that model performance was only slightly affected when only preoperative variables or only the most important predictive features were included.

Keywords: acute kidney injury, machine learning, noncardiac surgery, prediction model

[a]*Department of Anesthesiology and Pain Medicine, Hubei Key Laboratory of Geriatric Anesthesia and Perioperative Brain Health, and Wuhan Clinical Research Center for Geriatric Anesthesia,* [b]*Health Management Center, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei and* [c]*Yidu Cloud Technology Inc, Beijing, People's Republic of China*

Rao Sun and Shiyong Li contributed equally to this article.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding authors. Address: Department of Anesthesiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No.1095 Jiefang Avenue, Wuhan 430030, Hubei Province, People's Republic of China. Tel.: + 86 278 366 3423. E-mail: 397060616@qq.com (X. Li); Tel.: +86 278 366 3173. E-mail: alluo@hust.edu.cn (A. Luo), and Tel.: +86 278 366 5480. E-mail: zqzhouhustjmz@hust.edu.cn (Z. Zhou).*

International Journal of Surgery (2024) 110:2950–2962

Received 1 December 2023; Accepted 15 February 2024

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 4 March 2024

http://dx.doi.org/10.1097/JS9.0000000000001237

Introduction

Acute kidney injury (AKI) is a common surgical complication characterized by a rapid decline in renal function[1,2]. Patients with AKI are at an increased risk of developing chronic kidney disease and end-stage renal disease, which has been associated with an increased risk of morbidity, mortality, and financial burdens[3–5]. Identifying high-risk patients for postoperative AKI early can facilitate the development of preventive and therapeutic management strategies, and prediction models can be helpful in this regard.

Machine learning has emerged as a promising approach for developing prediction models in recent years[6,7]. Although machine learning-based models have been traditionally viewed as black boxes with limited interpretability, the development of techniques such as Shapley Additive Explanations (SHAP)[8] has made machine learning models more interpretable and increased their value in clinical settings. A growing number of studies have utilized machine learning to develop prediction models for post-operative complications[9–13], including AKI[11,14]. However, most studies for predicting AKI focus solely on cardiac surgery[15–18] and cannot be directly applied to patients undergoing noncardiac surgery. Additionally, prior machine learning-based AKI prediction models had several limitations, such as a

narrow focus on specific procedures and small sample sizes[19–21], as well as the absence of key variables[10,11]. Furthermore, previous models included numerous features[10,14], which increased their vulnerability to missing data.

In this study, we utilized machine learning algorithms to build prediction models for postoperative AKI in noncardiac surgery. Initially, we used all available variables, including preoperative and intraoperative variables, and then only preoperative variables. SHAP analysis was used to determine the feature importance of variables. Additionally, we constructed simplified models that included only the most important features.

HIGHLIGHTS

- Gradient boosting decision trees performed the best among the machine learning algorithms we evaluated.
- Age, surgical duration, and preoperative serum gamma-glutamyltransferase were the most important features for predicting postoperative acute kidney injury.
- Inclusion of only preoperative variables or the most important predictive features did not substantially affect the performance of the model.

Methods

The study was reported in accordance with the statement strengthening the reporting of cohort studies in surgery (STROCSS) criteria[22] (Supplemental Digital Content 1, http://links.lww.com/JS9/C27). We obtained ethical approval from the local Ethics Committee, and informed consent requirements were waived because this study is retrospective in nature. The study protocol has been registered on ClinicalTrials.

We included adult surgical patients (age $\geq$ 18 years) who had a serum creatinine measurement within 10 days before surgery and at least one measurement within 7 days after surgery. Eligible surgeries encompassed general, thoracic, orthopedic, obstetric, gynecology, and neurosurgery procedures lasting longer than 1 h. If a patient underwent multiple surgeries meeting the inclusion criteria during the study period, only the first surgery was considered in the analysis. Patients with concurrent cardiac, vascular, urological, or transplant surgeries, an American Society of Anesthesiologists (ASA) physical status V, or end-stage renal disease [i.e. a glomerular filtration rate (eGFR) of 15 ml/min/1.73 m² or receiving hemodialysis] were excluded.

Patients who underwent surgical procedures at our hospital between July 2018 and April 2022 were included as the discovery cohort, and those between May 2022 and October 2022 were included as the validation cohort.

Explanatory variables

A total of 87 features were identified as explanatory variables and classified into preoperative and intraoperative categories (Supplemental Table 1, Supplemental Digital Content 2, http://links.lww.com/JS9/C28). Preoperative variables included patient demographics (e.g. sex, age, BMI), comorbidities (e.g. hypertension, diabetes mellitus, coronary artery disease, renal insufficiency), medication history [e.g. use of beta blockers, angiotensin-converting enzyme inhibitors (ACEIs) or angiotensin receptor blockers (ARBs), lipid-lowering drugs, oral hypoglycemic agents, insulin], preoperative laboratory testing [e.g. hemoglobin, white blood cell (WBC) count, serum albumin and creatinine, and eGFR], preoperative vital signs [e.g. body temperature, respiratory rate, heart rate, diastolic blood pressure (DBP), systolic blood pressure (SBP)], and ASA physical status. Intraoperative data encompassed surgical characteristics (e.g. type and duration of surgery), anesthesia characteristics (e.g. type and duration of anesthesia), intraoperative drug administration (e.g. use of inhalational anesthetics, muscle relaxants, opioids, and vasopressors), intraoperative vital signs [e.g. heart rate, DBP, SBP, percutaneous oxygen saturation (SpO_2)], intraoperative transfusion [e.g. plasma, red blood cells (RBC), platelets, albumin], and colloid administration. The data elements were extracted from the patient's electronic health record and anesthesia record.

Outcome assessment

Postoperative AKI was defined based on the Kidney Disease Improving Global Outcomes (KDIGO) creatinine criteria[23]. Specifically, it was determined by a serum creatinine increase of 26.5 μmol/l (0.3 mg/dl) within 48 h or 1.5 times the baseline level within 7 days after surgery. The baseline was established using the most recent serum creatinine measurement prior to the surgical procedure. To evaluate AKI, we extracted the serum creatinine measurement taken within 7 days following surgery.

Data preprocessing

Among the 87 explanatory variables, 47 had missing data, with the missing rate ranging from 0.01 to 3.77%. To handle missing values, categorical variables (e.g. ASA physical status) were imputed using the mode value, while continuous variables were imputed using the median value. A Yeo-Johnson transformation was performed for highly skewed continuous variables [including laboratory test results such as WBC, neutrophil, eosinophil, and basophil counts, alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyltransferase (GGT), total bilirubin (TBIL), direct bilirubin (DBIL), blood glucose, international normalized ratio (INR), blood urea nitrogen, and creatinine; vital signs such as body temperature and SpO_2; duration of surgery, and duration of anesthesia]. The continuous variables were then zero-centered. Nonbinary categorical variables (including ASA physical status, type of surgery, plasma transfusion volume, and RBC transfusion volume) were encoded using one-hot encoding[24]. The details of data preprocessing are shown in Supplemental Table 1 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28).

Model construction and evaluation

The discovery cohort was divided into training and test sets using a 7:3 ratio. Two prediction models were constructed based on all available variables, including preoperative and intraoperative variables, or only on the preoperative variables. A variety of machine learning algorithms, including logistic regression, random forest, extreme gradient boosting (XGBoost), and gradient boosting decision tree (GBDT), were employed to construct the models. For each machine learning algorithm, grid search was used to determine the optimal hyperparameter values for the

highest area under the receiver operating characteristic curve (AUROC). The established models were evaluated in the test set and validated in the validation cohort. The discrimination performance of the models was evaluated by calculating AUROCs with 95% CIs. The higher the AUROC, the better the discrimination performance. Calibration curves were used to assess the model's calibration. The closer a calibration curve is to the diagonal line, the better the calibration performance. To evaluate model robustness, a sensitivity analysis was performed using only cases without missing values.

Model interpretation

We used the SHAP analysis to interpret model predictions. The SHAP value of each predictive variable was calculated to determine its feature importance, that is, the influence on a prediction in terms of direction and range. A positive SHAP value indicated that the corresponding feature contributed to a higher risk of AKI, whereas a negative SHAP value indicated that the corresponding feature led to a lower risk of AKI. The magnitude of SHAP values represented how much each feature contributed toward prediction performance. We utilize beeswarm plots to illustrate the ranking of important features, waterfall plots to illustrate how each feature contributes to individual predictions, and dependency plots to illustrate the relationship between features and outcomes.

Model simplification

The best machine learning algorithm's feature importance ranking result was used to select the top 10, top 20, top 30, and top 40 features for constructing the simplified prediction models.

Statistical analysis

In this study, continuous variables were not normally distributed, as determined by the Shapiro–Wilk test. Therefore, they were expressed as median [interquartile range (IQR)] and compared with the Mann–Whitney *U* test. Categorical variables were expressed as counts (percentages) and compared using the χ^2 or Fisher exact test. Spearman correlations were used in correlation analysis. All *P*-values were 2-tailed, and a *P*-value of <0.05 was considered statistically significant. The statistical analysis was conducted using Python 3.8.8.

Results

Study population characteristics

A total of 76 457 patients were included in the discovery cohort, as shown in Figure 1. These patients had a median age of 52 years (IQR: 42–61 years), with males accounting for 37.9% and most patients (67.5%) being at ASA II. Nonabdominal general surgery (26.8%) and gynecology surgery (26.7%) were the most common surgeries performed. The median duration of surgery was 2.5 h (IQR: 1.8–3.7 h), with AKI developing after surgery in 1179 patients (1.5%). Table 1 presents the characteristics of the discovery cohort, both globally and stratified by the presence of AKI, while Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) compares population characteristics between the training and test sets. For both the training and test sets, the proportion of patients in each year was comparable (Supplemental Table 3, Supplemental Digital Content 2, http://links.lww.com/JS9/C28). Supplemental Table 4 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) and Supplemental Table 5 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) show the characteristics of training and test sets, stratified by the presence of AKI. The validation cohort included a total of 11 910 patients, with similar population characteristics to those in the discovery cohort (Supplemental Table 6, Supplemental Digital Content 2, http://links.lww.com/JS9/C28).

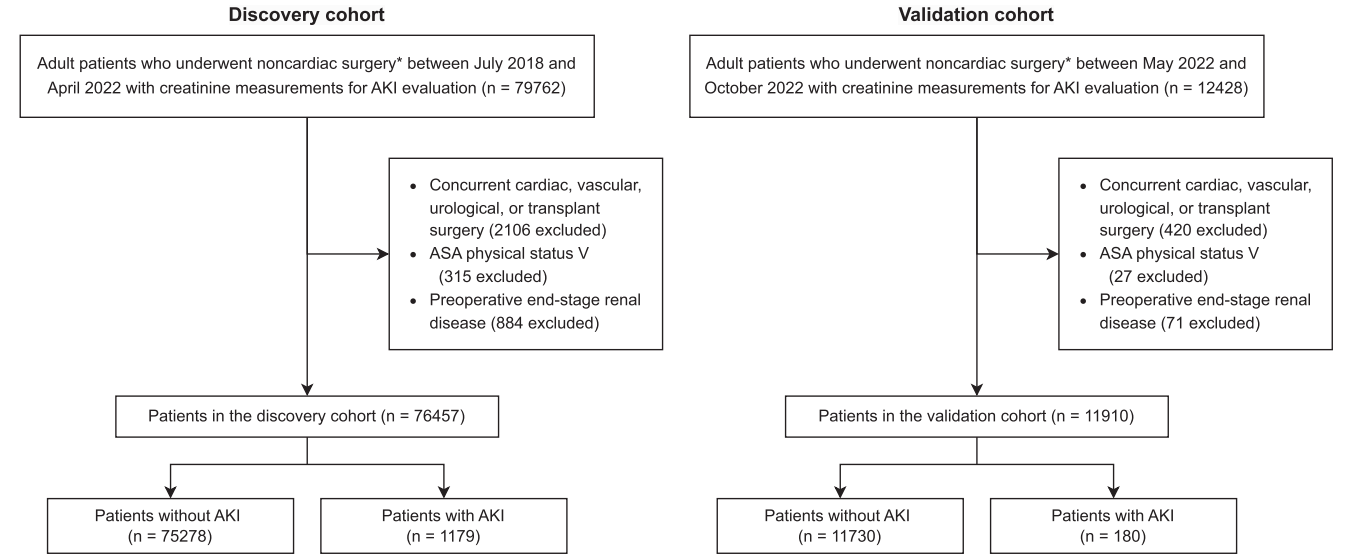

Figure 1. Flow diagram for patient selection. *Noncardiac surgeries encompassed general, thoracic, orthopedic, obstetric, gynecology, and neurosurgery procedures lasting longer than 1 h. AKI, acute kidney injury; ASA, American Society of Anesthesiologists.

Table 1

Characteristics of the discovery cohort.

	Overall ($N = 76\,457$)	No AKI ($n = 75\,278$)	AKI ($n = 1179$)	*P*
Male sex	28 986 (37.9%)	28 261 (37.5%)	725 (61.5%)	< 0.001
Age, year	52 (42–61)	52 (42–60)	61 (53–68)	< 0.001
BMI, kg/m^2	23.0 (20.8–25.2)	23.0 (20.8–25.2)	23.7 (21.3–26.0)	< 0.001
Any cancer	17 815 (23.3%)	17 415 (23.1%)	400 (33.9%)	< 0.001
Smoker	10 222 (13.4%)	9911 (13.2%)	311 (26.4%)	< 0.001
Alcohol use	6133 (8.0%)	5913 (7.9%)	220 (18.7%)	< 0.001
Comorbidity				
Hypertension	8500 (11.1%)	8201 (10.9%)	299 (25.4%)	< 0.001
Diabetes mellitus	3332 (4.4%)	3217 (4.3%)	115 (9.8%)	< 0.001
Coronary artery disease	1917 (2.5%)	1848 (2.5%)	69 (5.9%)	< 0.001
Hyperlipidemia	3788 (5.0%)	3713 (4.9%)	75 (6.4%)	0.030
Renal insufficiency	868 (1.1%)	734 (1.0%)	134 (11.4%)	< 0.001
Peripheral artery disease	558 (0.7%)	529 (0.7%)	29 (2.5%)	< 0.001
Cirrhosis	886 (1.2%)	837 (1.1%)	49 (4.2%)	< 0.001
Anemia	29 433 (38.5%)	28 826 (38.3%)	607 (51.5%)	< 0.001
Preoperative medication				
Beta blocker	2793 (3.7%)	2707 (3.6%)	86 (7.3%)	< 0.001
ACEIs/ARBs	2333 (3.1%)	2245 (3.0%)	88 (7.5%)	< 0.001
Lipid-lowering drugs	1781 (2.3%)	1720 (2.3%)	61 (5.2%)	< 0.001
Oral hypoglycemic agents	1738 (2.3%)	1695 (2.3%)	43 (3.6%)	0.002
Insulin	14 902 (19.5%)	14 581 (19.4%)	321 (27.2%)	< 0.001
Preoperative laboratory testing				
Hemoglobin, g/l	127 (116–138)	127 (116–138)	125 (109–137)	< 0.001
Hematocrit, %	38.1 (35.0–41.2)	38.1 (35.0–41.2)	37.4 (33.2–40.9)	< 0.001
WBC count, $\times 10^9$/l	5.6 (4.6–6.9)	5.6 (4.6–6.9)	6.0 (4.8–7.6)	< 0.001
RBC count, $\times 10^{12}$/l	4.2 (3.9–4.6)	4.2 (3.9–4.6)	4.1 (3.7–4.5)	< 0.001
Neutrophil count, $\times 10^9$/l	3.2 (2.5–4.3)	3.2 (2.5–4.3)	3.7 (2.7–5.2)	< 0.001
Lymphocyte count, $\times 10^9$/l	1.6 (1.3–2.0)	1.6 (1.3–2.0)	1.4 (1.1–1.9)	< 0.001
Eosinophil count, $\times 10^9$/l	0.10 (0.06–0.17)	0.10 (0.06–0.17)	0.11 (0.06–0.20)	0.005
Basophil count, $\times 10^9$/l	0.02 (0.01–0.03)	0.02 (0.01–0.03)	0.02 (0.01–0.03)	0.001
Platelet count, $\times 10^9$/l	220 (178–268)	220 (178–268)	205 (155–260)	< 0.001
Total protein, g/l	68.7 (65.2–72.6)	68.7 (65.2–72.6)	67.8 (64.0–71.9)	< 0.001
Albumin, g/l	41.1 (38.7–43.6)	41.1 (38.7–43.6)	39.3 (36.1–42.0)	< 0.001
Globulin, g/l	27.6 (25.1–30.5)	27.6 (25.1–30.5)	28.7 (25.7–32.0)	< 0.001
ALT, U/l	14 (10–22)	14 (10–22)	17 (11–26)	< 0.001
AST, U/l	18 (15–22)	18 (15–22)	20 (15–28)	< 0.001
GGT, U/l	19 (13–32)	19 (13–32)	27 (18–53)	< 0.001
TBIL, umol/l	8.9 (6.5–12.4)	8.9 (6.5–12.4)	9.1 (6.5–13.1)	0.032
DBIL, umol/l	3.5 (2.7–4.7)	3.5 (2.7–4.7)	3.8 (2.8–5.3)	< 0.001
Total cholesterol, mmol/l	4.0 (3.5–4.6)	4.0 (3.5–4.6)	3.9 (3.2–4.6)	< 0.001
Glucose, mmol/l	5.0 (4.7–5.6)	5.0 (4.7–5.5)	5.3 (4.9–6.4)	< 0.001
TT, s	16.7 (16.0–17.3)	16.7 (16.0–17.3)	16.9 (16.1–17.6)	< 0.001
PT, s	13.2 (12.7–13.7)	13.2 (12.7–13.7)	13.3 (12.8–14.0)	< 0.001
APTT, s	37.1 (34.8–39.7)	37.1 (34.8–39.7)	37.6 (34.9–40.8)	< 0.001
PTA, %	100 (92–108)	100 (92–108)	97 (87–107)	< 0.001
INR	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.0 (1.0–1.1)	< 0.001
Potassium, mmol/l	4.0 (3.8–4.2)	4.0 (3.8–4.2)	4.1 (3.8–4.3)	0.035
Sodium, mmol/l	140 (139–142)	140 (139–142)	140 (139–142)	0.039
Chloride, mmol/l	104 (102–105)	104 (102–105)	104 (101–105)	0.591
Blood urea nitrogen, mmol/l	4.8 (3.9–5.8)	4.7 (3.9–5.8)	5.3 (4.2–6.8)	< 0.001
Uric acid, umol/l	288 (237–350)	288 (237–350)	318 (250–392)	< 0.001
Creatinine, umol/l	63 (55–75)	63 (55–75)	74 (57–93)	< 0.001
eGFR, ml/min/1.73 m^2	101.0 (90.5–111.0)	101.1 (90.8–111.0)	90.9 (69.6–102.9)	< 0.001
Preoperative vital sign				
Body temperature, °C	36.5 (36.3–36.6)	36.5 (36.3–36.6)	36.5 (36.3–36.7)	0.102
Respiratory rate, bpm	20 (19–20)	20 (19–20)	20 (19–20)	0.001
Pulse rate, bpm	78 (73–80)	78 (73–80)	78 (74–82)	0.059
DBP, mmHg	76 (70–84)	76 (70–84)	78 (71–85)	< 0.001
SBP, mmHg	121 (111–132)	121 (111–132)	128 (118–140)	< 0.001
ASA physical status				< 0.001
I	11 640 (15.2%)	11 590 (15.4%)	50 (4.2%)	
II	51 627 (67.5%)	51 040 (67.8%)	587 (49.8%)	

Table 1

(*Continued*)

	Overall ($N=76\,457$)	No AKI ($n=75\,278$)	AKI ($n=1179$)	*P*
III	12 847 (16.8%)	12 347 (16.4%)	500 (42.4%)	
IV	343 (0.4%)	301 (0.4%)	42 (3.6%)	
Type of surgery				< 0.001
Obstetric	480 (0.6%)	464 (0.6%)	16 (1.4%)	
Gynecology	20 422 (26.7%)	20 299 (27.0%)	123 (10.4%)	
General (nonabdominal)	20 514 (26.8%)	20 064 (26.7%)	450 (38.2%)	
General (intra-abdominal)	2474 (3.2%)	2454 (3.3%)	20 (1.7%)	
Neurosurgery	5889 (7.7%)	5787 (7.7%)	102 (8.7%)	
Thoracic	15 650 (20.5%)	15 341 (20.4%)	309 (26.2%)	
Orthopedic	11 028 (14.4%)	10 869 (14.4%)	159 (13.5%)	
Emergency surgery	12 600 (16.5%)	12 382 (16.4%)	218 (18.5%)	0.066
Duration of surgery, h	2.5 (1.8–3.7)	2.5 (1.8–3.6)	3.6 (2.5–5.2)	< 0.001
Duration of anesthesia, h	3.2 (2.3–4.4)	3.2 (2.3–4.4)	4.4 (3.2–6.0)	< 0.001
Intraoperative drug administration				
Inhalational anesthetics	73 567 (96.2%)	72 429 (96.2%)	1138 (96.5%)	0.637
Muscle relaxants	73 957 (96.7%)	72 810 (96.7%)	1147 (97.3%)	0.318
Opioids	75 870 (99.2%)	74 696 (99.2%)	1174 (99.6%)	0.232
Intravenous anesthetics[a]	75 885 (99.3%)	74 719 (99.3%)	1166 (98.9%)	0.210
Local anesthetics	47 814 (62.5%)	46 993 (62.4%)	821 (69.6%)	< 0.001
Vasopressors	48 137 (63.0%)	47 172 (62.7%)	965 (81.8%)	< 0.001
Type of anesthesia				0.249
General anesthesia	74 247 (97.1%)	73 095 (97.1%)	1152 (97.7%)	
Nongeneral anesthesia	2210 (2.9%)	2183 (2.9%)	27 (2.3%)	
Intraoperative vital sign				
Maximum heart rate, bpm	76 (70–84)	76 (70–84)	79 (71–89)	< 0.001
Maximum DBP, mmHg	78 (72–85)	78 (72–85)	80 (75–88)	< 0.001
Maximum SBP, mmHg	129 (120–141)	128 (120–141)	136 (125–150)	< 0.001
Maximum SpO_2, %	100 (99–100)	100 (99–100)	100 (99–100)	0.269
Minimum heart rate, bpm	68 (62–75)	68 (62–75)	69 (63–77)	< 0.001
Minimum DBP, mmHg	65 (57–70)	65 (57–70)	62 (54–70)	< 0.001
Minimum SBP, mmHg	110 (100–115)	110 (100–115)	108 (95–116)	< 0.001
Minimum SpO_2, %	99 (98–100)	99 (98–100)	99 (98–100)	< 0.001
Percentage change in intraoperative SBP from baseline (minimum), %	7.8 (– 1.8–18.7)	7.8 (– 1.8–18.7)	8.0 (– 3.4–19.8)	0.602
Percentage change in intraoperative SBP from baseline (maximum), %	– 10.8 (– 20.5, – 0.8)	– 10.7 (– 20.3, 0.0)	– 16.7 (– 27.4, – 6.6)	< 0.001
Percentage change in intraoperative DBP from baseline (minimum), %	2.7 (– 7.6–14.7)	2.7 (– 7.6–14.7)	4.2 (– 6.0–15.9)	0.002
Percentage change in intraoperative DBP from baseline (maximum), %	– 16.7 (– 27.1, – 5.7)	– 16.7 (– 27.0, – 5.6)	– 21.1 (– 31.8, – 9.2)	< 0.001
Intraoperative transfusion				
Blood transfusion	8145 (10.7%)	7787 (10.3%)	358 (30.4%)	< 0.001
Plasma transfusion				< 0.001
0 ml	71 026 (92.9%)	70 111 (93.1%)	915 (77.6%)	
0–200 ml	1652 (2.2%)	1601 (2.1%)	51 (4.3%)	
200–400 ml	2568 (3.4%)	2451 (3.3%)	117 (9.9%)	
> 400 ml	1211 (1.6%)	1115 (1.5%)	96 (8.1%)	
RBC transfusion				< 0.001
0 ml	69 020 (90.3%)	68 171 (90.6%)	849 (72.0%)	
0–400 ml	3245 (4.2%)	3140 (4.2%)	105 (8.9%)	
400–800 ml	3033 (4.0%)	2901 (3.9%)	132 (11.2%)	
> 800 ml	1159 (1.5%)	1066 (1.4%)	93 (7.9%)	
Platelet transfusion	213 (0.3%)	188 (0.2%)	25 (2.1%)	< 0.001
Autologous blood transfusion	1713 (2.2%)	1642 (2.2%)	71 (6.0%)	< 0.001
Cryoprecipitate transfusion	221 (0.3%)	194 (0.3%)	27 (2.3%)	< 0.001
Albumin transfusion	6119 (8.0%)	5946 (7.9%)	173 (14.7%)	< 0.001
Intraoperative colloid administration	6737 (8.8%)	6557 (8.7%)	180 (15.3%)	< 0.001

Values are expressed as median (interquartile range) or number of patients (%).

[a]The term 'intravenous anesthetics' refers to any dose of intravenous anesthetic administered with or without inhalational anesthetics.

ACEIs, angiotensin-converting enzyme inhibitors; AKI, acute kidney injury; ALT, alanine transaminase; APTT, activated partial thromboplastin time; ARBs, angiotensin receptor blockers; ASA, American Society of Anesthesiologists; AST, aspartate aminotransferase; DBIL, direct bilirubin; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; GGT, gamma-glutamyltransferase; INR, international normalized ratio; PT, prothrombin time; PTA, prothrombin activity; RBC, red blood cell; SBP, systolic blood pressure; SpO_2, percutaneous oxygen saturation; TBIL, total bilirubin; TT, thrombin time; WBC, white blood cell.

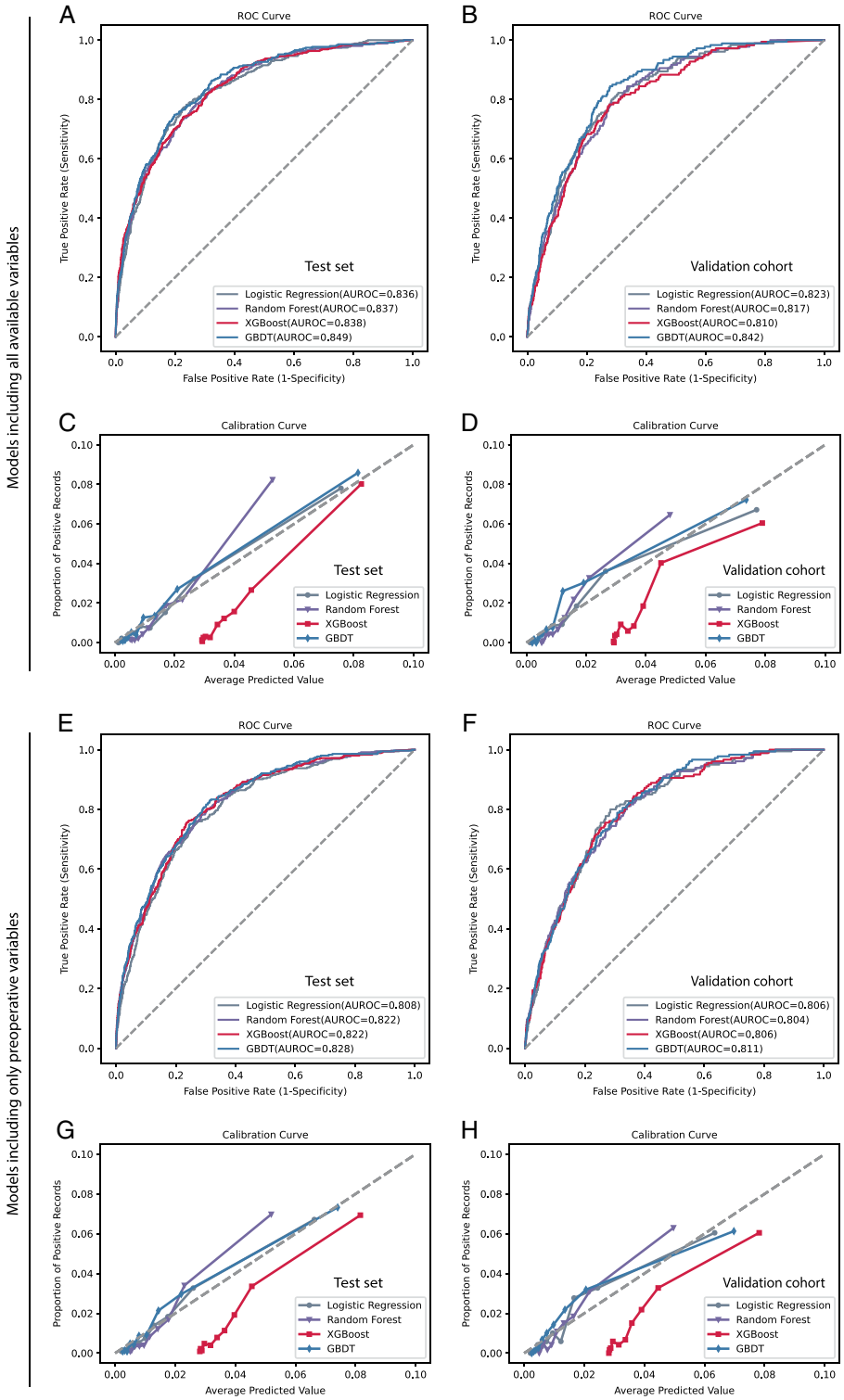

Figure 2. ROC curves and calibration curves for machine learning models in the test set and validation cohort. Models were constructed using all available variables (A–D) or only preoperative variables (E–H). The area under the AUROC was used as a measure of model discrimination (A, B, E, F). The higher the AUROC, the better the discrimination performance. Calibration curves (C, D, G, H) were used to assess the model's calibration. The closer a calibration curve is to the diagonal line, the better the calibration performance. AUROC, area under the receiver operating characteristic curve; GBDT, gradient boosting decision tree; ROC, receiver operating characteristic; XGBoost, extreme gradient boosting.

Model performance

First, we built prediction models using all available variables. As shown in Figure 2A and Table 2, the GBDT algorithm delivered the best discrimination performance in terms of AUROC [0.849 (95% CI: 0.835–0.863)], followed by XGBoost [0.838 (95% CI: 0.823–0.852)], random forest [0.837 (95% CI: 0.822–0.851)], and logistic regression [0.836 (95% CI: 0.821–0.851)] in the test set. There was a similar pattern in the validation cohort, with the GBDT algorithm exhibiting the highest AUROC (Fig. 2B and Table 2). Calibration plots indicate that GBDT and logistic regression models performed better in calibration, as their curves aligned more closely with the diagonal line (Fig. 2C and D).

Next, we built prediction models using only preoperative variables. In the test set, the GBDT algorithm still delivered the best discrimination performance in terms of AUROC [0.828 (95% CI: 0.813–0.843)], followed by XGBoost [0.822 (95% CI: 0.806–0.838)], random forest [0.822 (95% CI: 0.806–0.838)], and logistic regression [0.808 (95% CI: 0.791–0.825)] (Fig. 2E and Table 2). The GBDT algorithm had the highest AUROC in the validation cohort as well (Fig. 2F and Table 2). Calibration plots indicate that the models constructed using the GBDT algorithm and logistic regression had better calibration performance (Fig. 2G and H).

A sensitivity analysis was then performed to evaluate model robustness using only cases without missing values. The results indicated that the GBDT algorithm still had the highest AUROC (Table 3).

Model interpretation

For the prediction models developed by GBDT algorithm, we used the SHAP analysis to quantify the influence of different features on the model predictions. Figure 3A and Supplemental Figure 1 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) show the beeswarm plots for the top 20 and top 40 features in the model constructed by all available variables, among which age, duration of surgery, preoperative serum GGT and creatinine, as well as ASA physical status III rank among the top five. Figure 3B and C show the waterfall plots illustrating how each feature contributes to individual predictions. To visualize the relationship between features and outcomes, the SHAP dependence plots of the top 20 most important features are shown in Figure 4 (11 continuous variables) and Supplemental Figure 2 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) (nine categorical variables). The results show that older age, longer duration of surgery, higher preoperative GGT level, higher BMI, extremely high intraoperative SBP, and an extreme drop in intraoperative SBP from baseline, were related to increased risk of postoperative AKI. A U-shaped relationship was observed between preoperative creatinine, SBP, glucose, and TBIL and postoperative AKI risk. This indicated that both too low and too high levels of these parameters were associated with an increased risk of AKI. Moreover, we found that male sex, ASA physical status III, pre-existing hypertension, cancer, alcohol use before surgery, having thoracic surgery, as well as intraoperative blood transfusion were associated with a higher risk of postoperative AKI.

Figure 5A and Supplemental Figure 3 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) show the beeswarm plots of the top 20 and top 40 features in the model constructed only by preoperative variables, with age, preoperative serum GGT and creatinine, ASA physical status III, and preoperative SBP ranking as the top five features. Figure 5B and C show the waterfall plots illustrating how each feature contributes to individual predictions. We also draw SHAP dependence plots to visualize the relationship between the top 20 features and outcomes. As shown in the Supplemental Figure 4 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28), older age, higher preoperative GGT level, higher BMI, longer APTT, as well as lower preoperative serum albumin level, RBC, lymphocyte, and platelet counts are associated with a higher risk of AKI. We observed a U-shaped relationship between preoperative creatinine, glucose, uric acid and TBIL, and postoperative AKI risk.

The potential interactions among the top 10 features are shown in Supplemental Figures 5 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28) and 6 (Supplemental Digital Content 2, http://links.lww.com/JS9/C28). As shown in these figures, there may be some interactions between important features, such as age and preoperative creatinine levels and ASA physical status. Older patients (> 60 years old) have higher levels of creatinine and poor ASA physical status than younger patients.

Model simplification

To simplify the prediction models, we constructed models based solely on the top 40, top 30, top 20, and top 10 features. The results showed that when the features were gradually reduced, the AUROCs decreased slightly. When all available variables were used in prediction models, the AUROCs decreased from 0.852 (including the top 40 features) to 0.839 (including the top 10 features) in the test set (Table 4). When only preoperative variables were used in prediction models, the AUROCs decreased from 0.830 (including the top 40 features) to 0.818 (including the top 10 features) in the test set (Table 4). We also observed a similar pattern in the validation cohort (Table 4).

Table 2

AUROCs of machine learning-based models trained using imputed data.

	Preoperative and intraoperative variables		Preoperative variables	
Machine learning algorithm	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort
Logistic regression	0.836 (0.821–0.851)	0.823 (0.801–0.845)	0.808 (0.791–0.825)	0.806 (0.782–0.830)
Random forest	0.837 (0.822–0.851)	0.817 (0.795–0.840)	0.822 (0.806–0.838)	0.804 (0.780–0.828)
XGBoost	0.838 (0.823–0.852)	0.810 (0.787–0.833)	0.822 (0.806–0.838)	0.806 (0.782–0.829)
GBDT	0.849 (0.835–0.863)	0.842 (0.821–0.862)	0.828 (0.813–0.843)	0.811 (0.787–0.834)

AUROC, area under the receiver operating characteristic curve; GBDT, gradient boosting decision tree; XGBoost, extreme gradient boosting.

Table 3

AUROCs of machine learning models trained using data without missing values.

Machine learning algorithm	All variables		Preoperative variables	
	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort
Logistic regression	0.844 (0.829–0.859)	0.818 (0.794–0.842)	0.808 (0.791–0.825)	0.811 (0.786–0.835)
Random forest	0.837 (0.821–0.853)	0.813 (0.789–0.837)	0.822 (0.806–0.838)	0.803 (0.778–0.828)
XGBoost	0.839 (0.823–0.855)	0.824 (0.801–0.847)	0.817 (0.801–0.833)	0.792 (0.766–0.818)
GBDT	0.853 (0.838–0.867)	0.830 (0.807–0.852)	0.828 (0.813–0.843)	0.813 (0.789–0.837)

AUROC, area under the receiver operating characteristic curve; GBDT, gradient boosting decision tree; XGBoost, extreme gradient boosting.

Discussion

In this study, 76 457 patients were enrolled to construct prediction models for postoperative AKI. These models were subsequently prospectively validated on 11 910 patients in a separate validation cohort. Our results showed that GBDT demonstrated the best performance among the considered machine learning algorithms. The prediction model that uses only preoperative variables had lower predictive performance than the model that incorporates both preoperative and intraoperative variables, but it still shows good predictive capability. As for the importance of the predictive variables, age, preoperative serum GGT and creatinine levels, ASA physical status, and preoperative blood pressure were the most important preoperative variables, whereas duration and type of surgery, intraoperative blood pressure, and intraoperative vasopressor use were the most important intraoperative variables. We also found that model performance only decreased slightly when gradually reducing the predictive features and including only the most important ones.

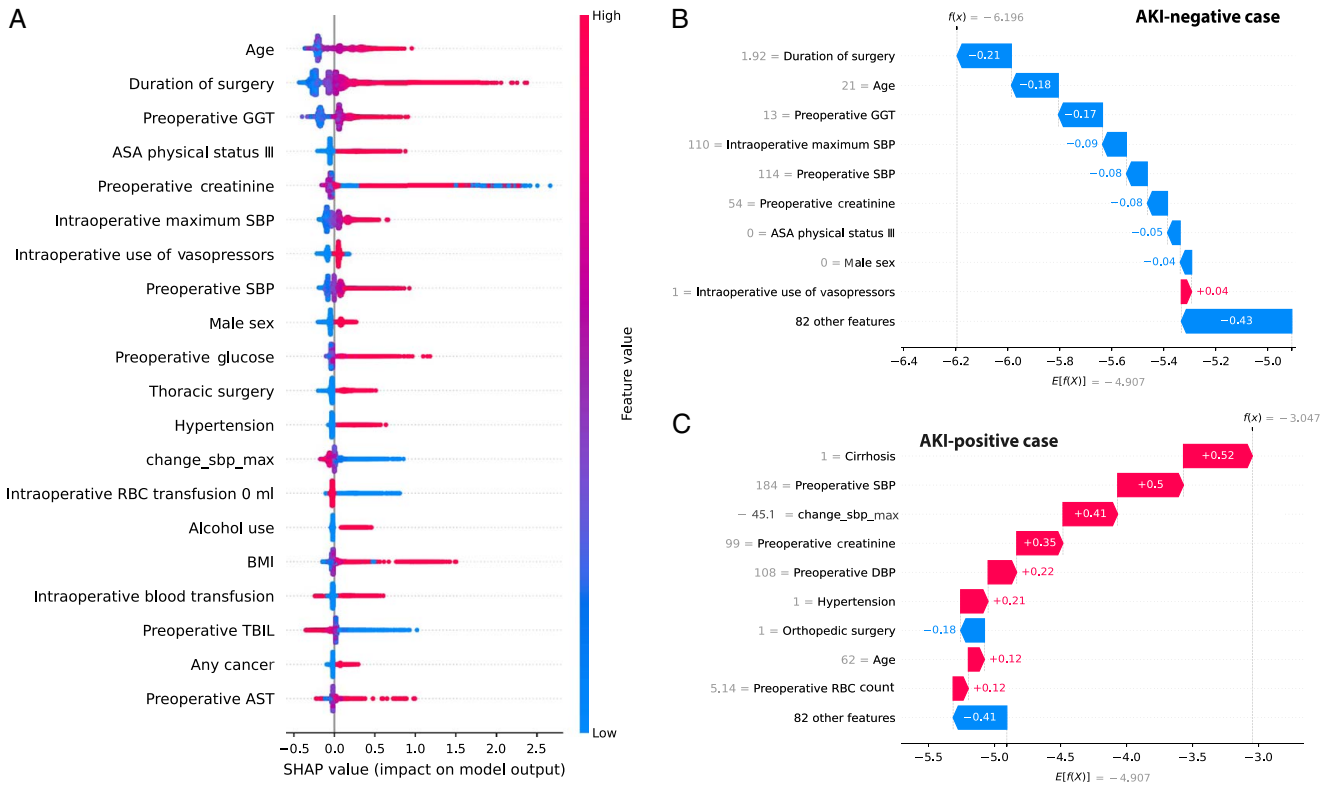

Figure 3. Interpretation of the model constructed by the GBDT algorithm using all available variables. (A) The beeswarm plot for the top 20 features in the model. A dot is created for each patient in each feature, with red denoting a higher feature value and blue denoting a lower feature value. The *x*-axis represents the SHAP values that describe the impact of each feature on model prediction. Positive SHAP values indicate an increased risk of postoperative AKI, whereas negative SHAP values indicate a decreased risk. The features are sorted by the sum of the SHAP value magnitudes. (B, C) The waterfall plots illustrating how each feature contributes to individual predictions (B is an AKI-negative case; C is an AKI-positive case). On a waterfall plot, the value at the bottom represents the expected value of the model output, and each row represents the contribution of each feature to the model output. A red arrow indicates an increased risk of postoperative AKI, while a blue arrow indicates a decreased risk. The gray text before the feature names shows the value of each feature for the case. AKI, acute kidney injury; AST, aspartate aminotransferase; ASA, American Society of Anesthesiologists; DBP, diastolic blood pressure; GGT, gamma-alutamyltransferase; RBC, red blood cell; SHAP, Shapley additive explanations; SBP, systolic blood pressure; TBIL, total bilirubin; change_sbp_max, percentage change in intraoperative SBP from baseline (maximum).

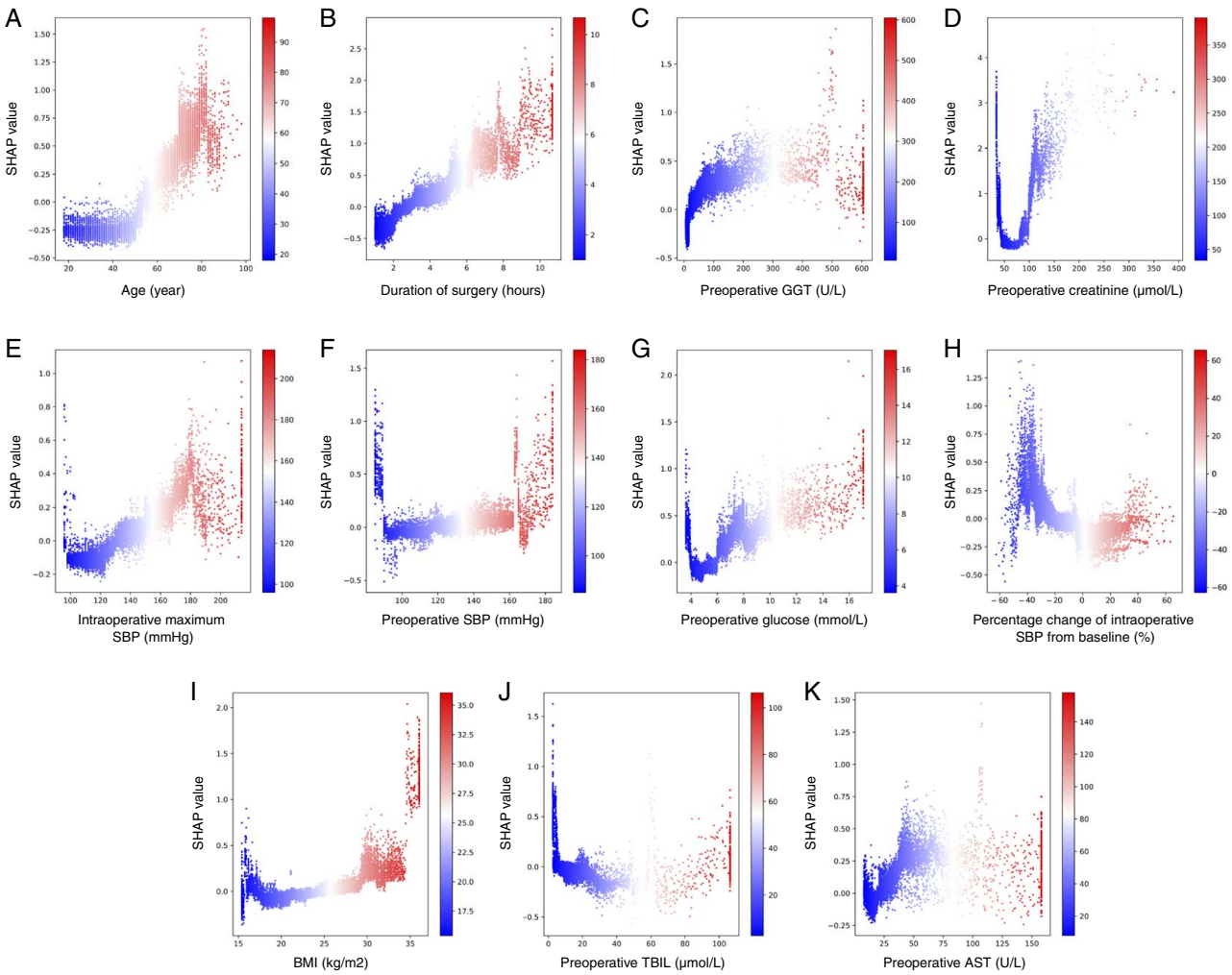

Figure 4. SHAP dependence plots for the continuous variables among the top 20 features in the GBDT algorithm model using all available variables. (A-K) The SHAP dependence plots for each variable Each dot represents a single prediction. The *x*-axis represents the actual values of features, while the *y*-axis represents the SHAP values. Feature values are also indicated by color bars, the redder the color, the higher the value. Positive SHAP values indicate an increased risk of postoperative AKI, whereas negative SHAP values indicate a decreased risk. AST, aspartate aminotransferase; GGT, gamma-alutamyltransferase; SHAP, Shapley additive explanations; SBP, systolic blood pressure; TBIL, total bilirubin.

Several studies with large sample sizes have utilized machine learning algorithms to develop prediction models for postoperative AKI. Hofer *et al.*[11] constructed an AKI prediction model using the deep neural network algorithm. The model contains 52 features and had an AUROC of 0.792 (95% CI: 0.775–0.808). Although the model includes detailed data on intraoperative vital signs and medication, some variables such as the duration of surgery and preoperative laboratory tests were not considered, which limits the model's ability to predict outcomes. MySurgeryRisk is an automated machine learning model developed by Bihorac *et al.*[25] in 2019. The model is designed to predict postoperative complications, including AKI. In a subsequent study[9], the authors conducted prospective internal validation and found that the AUROC of the model, which included all 135 features, was 0.84 (95% CI: 0.83–0.84), and that AUROC after streamlining the model to 55 features was 0.82 (95% CI: 0.82–0.83). The model was embedded into the medical system for the prediction of complications, and it was found to be highly consistent with the predictions of physicians' judgments. It is a meaningful practice that confirms the effectiveness of machine learning-based prediction models in clinical settings. The machine learning-based AKI prediction model developed by Lei *et al.*[14] divided the predictive variables into preoperative and perioperative categories. In their study, they found that when incorporating only preoperative variables (338 variables), the AUROC was 0.804 (95% CI: 0.788–0.819), and when adding intraoperative variables (339 variables), the AUROC was only modestly increased to 0.817 (95% CI: 0.802–0.832). Xue *et al.*[10] developed another important machine learning-based prediction model for AKI. They found that the GBDT algorithm had the highest AUROC of 0.848 (95% CI: 0.846–0.851). Although they included a total of 711 variables to build prediction model, some key variables, such as duration of surgery, intraoperative blood transfusion data, and vasopressor utilization, were not accounted for in the model. These variables have been proven to be risk factors for AKI and have been considered important features in previous models as well as our model.

Although machine learning has the major advantage of using a large amount of information to make predictions automatically, simplification of the model remains important since more features

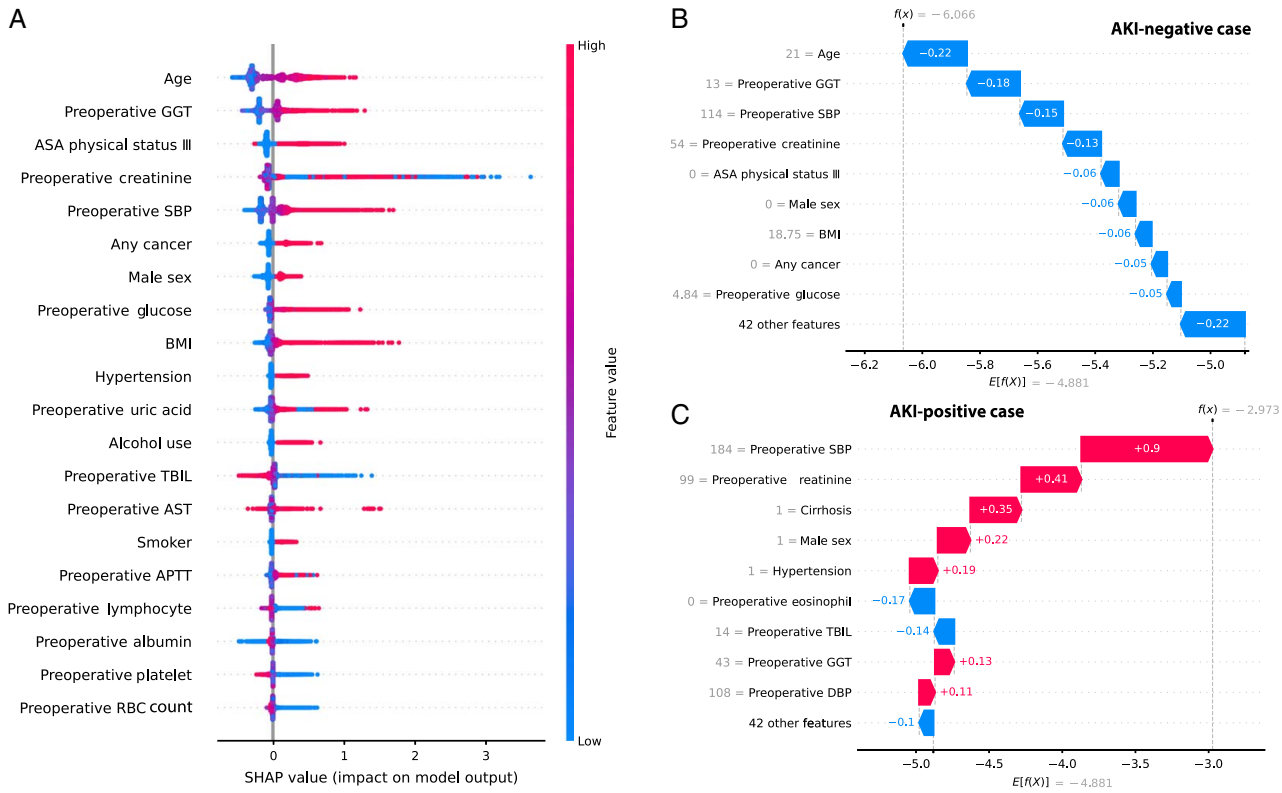

Figure 5. Interpretation of the model constructed by the GBDT algorithm using only preoperative variables. (A) The beeswarm plot for the top 20 features in the model. A dot is created for each patient in each feature, with red denoting a higher feature value and blue denoting a lower feature value. The *x*-axis represents the SHAP values that describe the impact of each feature on model prediction. Positive SHAP values indicate an increased risk of postoperative AKI, whereas negative SHAP values indicate a decreased risk. The features are sorted by the sum of the SHAP value magnitudes. (B, C) The waterfall plots illustrating how each feature contributes to individual predictions (B is an AKI-negative case; C is an AKI-positive case). On a waterfall plot, the value at the bottom represents the expected value of the model output, and each row represents the contribution of each feature to the model output. A red arrow indicates an increased risk of postoperative AKI, while a blue arrow indicates a decreased risk. The gray text before the feature names shows the value of each feature for the case. ASA, American Society of Anesthesiologists; AKI, acute kidney injury; APTT, activated partial thromboplastin time; AST, aspartate aminotransferase; DBP, diastolic blood pressure; GGT, gamma-alutamyltransferase; RBC, red blood cell; SBP, systolic blood pressure; SHAP, Shapley additive explanations; TBIL, total bilirubin.

required increase the chance that it will be impacted by missing data. In this study, we first built prediction models using all variables and then only preoperative variables. The model based solely on preoperative variables had lower predictive accuracy than the model that included all variables (AUROCs: 0.828 vs. 0.849), but it still had good predictive performance. We further simplified the models by selecting the top 10, top 20, top 30, and top 40 features to build a simplified prediction model. Our results showed that even though only 10 features were included, the model still performed well with AUROCs of 0.839 and 0.818.

Our prediction models found that older age and male sex were associated with a higher risk of AKI. This is consistent with many other studies[26–28]. Another important predictor of post-operative AKI is the ASA. An elevated ASA indicates multiple and serious comorbidities, which are strongly associated with an increased risk of postoperative AKI[26]. Our model identified hypertension as a significant comorbidity predictor. This is in agreement with previous studies that have demonstrated the relationship between hypertension and AKI[29,30].

Our study showed that higher preoperative serum GGT level, as well as lower preoperative serum albumin level, RBC, lymphocyte, and platelet counts are associated with a higher risk of AKI. While GGT is typically used to assess liver function, it has been linked to vascular endothelial dysfunction, causing urinary albumin extraction and kidney damage[31,32], and has emerged as a biomarker for predicting contrast-induced nephropathy[33]. Serum albumin is essential for maintaining colloid osmotic pressure, enhancing effective circulation volume, promoting renal blood flow, and conserving renal function[34]. In both cardiac and non-cardiac surgeries, hypoalbuminemia is a known risk factor for postoperative AKI[35,36]. In regard to blood cells, previous studies have indicated that low preoperative lymphocyte and platelet counts are associated with an increased risk of postoperative AKI[37,38]. There is also a link between low RBC count and the development of AKI[26]. A reduced hemoglobin level results in a reduced oxygen carrying capacity and increases the risk of cellular damage[39].

We observed a U-shaped relationship between preoperative creatinine, uric acid, glucose, TBIL, and postoperative AKI risk. This indicated that both too low and too high levels of these parameters were associated with an increased risk of AKI. Generally, serum creatinine can be used as an indicator of renal function. An elevated serum creatinine level indicates impaired baseline renal function, which increases postoperative AKI risk[28,40]. However, serum creatinine levels may underestimate renal function impairment in populations with diminished muscle mass[41], such as those suffering from sarcopenia or malnutrition. It has been demonstrated that these populations are at an

Table 4

AUROCs of models with various features.

Features included in the model	All variables		Preoperative variables	
	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort	AUROC (95% CI) in the test set	AUROC (95% CI) in the validation cohort
Top 40	0.852 (0.838–0.865)	0.840 (0.820–0.860)	0.830 (0.815–0.845)	0.814 (0.791–0.837)
Top 30	0.849 (0.835–0.863)	0.833 (0.812–0.854)	0.825 (0.810–0.841)	0.809 (0.786–0.833)
Top 20	0.848 (0.834–0.862)	0.827 (0.806–0.849)	0.826 (0.810–0.842)	0.807 (0.784–0.831)
Top 10	0.839 (0.824–0.853)	0.807 (0.783–0.83)	0.818 (0.802–0.834)	0.800 (0.776–0.825)

The models were developed using the gradient boosting decision tree (GBDT) algorithm.
AUROC, area under the receiver operating characteristic curve.

increased risk of AKI[2,42]. In a recent study, a U-shaped relationship has been demonstrated between preoperative creatinine and postoperative AKI[43]. Similarly, both decreased and elevated serum uric acid levels were associated with a reduction in kidney function[44]. As well as hyperuricemia being a direct risk factor for kidney failure, hypouricemia has also been found to worsen renal function[45]. Elevated blood glucose levels can increase the risk of AKI[46], and it has been demonstrated that intensive glucose control during the perioperative period can reduce the risk of postoperative AKI[47]. Meanwhile, severe hypoglycemia was also reported to be associated with renal function decline[48,49]. Bilirubin can bind to albumin and exhibit antioxidative and anti-inflammatory properties, thereby protecting the kidneys[50,51]. However, this protective effect is based on mild elevations of bilirubin levels within normal ranges (< 1.2 mg/dl), and severe hyperbilirubinemia (total bilirubin > 2.0 mg/dl) is a risk factor for AKI[52]. Overall, given the U-shaped relationship between these features and postoperative AKI, future studies may consider converting these variables into categorical data with appropriate cutoff points and incorporating them into prediction models. However, these U-shaped relationships were based on our post-hoc analysis and should be confirmed by further research.

Our study found that the length and type of surgery, as well as intraoperative hemodynamics, vasopressors use, and blood transfusions were important predictors of postoperative AKI. In line with previous research[53], we found that different types of surgery are associated with different incidence of postoperative AKI. Although the type of surgery is not modifiable, shortening the surgery time may reduce the risk of AKI for patients, especially high-risk patients[26]. Regarding hemodynamics, we found that extremely high intraoperative SBP and an extreme drop in intraoperative SBP from baseline were associated with increased risk of AKI. This finding highlights the importance of goal-directed hemodynamic therapy and maintaining euvolaemia during surgery to prevent AKI[27]. We also found that use of vasopressors was associated with increased risk of AKI. This can be explained that use of vasopressors represents a surrogate marker for less stable hemodynamics[53]. Moreover, there are also data showing that higher vasopressor use affects renal function due to decreased renal perfusion[54]. In regard to blood transfusion, preoperative anemia is a risk factor for postoperative AKI, but blood transfusion itself increases the risk of postoperative AKI[55]. Therefore, individualized blood transfusion plans need to be developed based on a patient's condition[56,57].

According to our study, the prediction model based solely on preoperative variables showed a reliable predictive ability. Thus, the model can be utilized to identify high-risk patients prior to surgery and develop perioperative care management plans at an early stage in such patients to prevent AKI. According to the Acute Disease Quality Initiative and the Perioperative Quality Initiative, a variety of strategies have been suggested to reduce the risk of postoperative AKI. These include discontinuing ACEIs and ARBs prior to surgery, using goal-directed haemodynamic therapy during surgery, maintaining an intraoperative mean arterial blood pressure > 65 mmHg, using balanced crystalloids instead of 0.9% saline, maintaining euvolaemia, as well as treating hypotension and hyperglycemia following surgery.

This study has several strengths. First, the model was trained using a dataset with a relatively large sample size and a lower proportion of missing data. Second, we used sensitivity analysis to evaluate the robustness of the model and prospectively validate it. Last, we evaluated the importance of the model's features and simplified it by incorporating important features, and we found that the simplified models also performed well.

This study also has some limitations. First, the incidence of AKI in our study was lower than that in other studies. It may be due to the fact that we excluded surgeries which have a high incidence of AKI, such as those involving cardiac, vascular, urological, or transplant surgery. Second, we did not use time-series data when analyzing the intraoperative vital signs, but rather only the highest and lowest intraoperative values and their percentage changes from baseline. Finally, although we performed prospective internal validation of our model, external applicability requires further testing since the model came from a single center.

Overall, the machine learning models we developed had satisfactory predictive performance for identifying high-risk postoperative AKI patients. We also found that model performance only decreased slightly when including only preoperative variables or only the most important predictive features.

Ethical approval

This study was approved by the Ethics Committee of Tongji Hospital, Tongji Medical College, Huazhong University of Science (TJ-IRB20230741).

Consent

The requirement for informed consent was waived due to the retrospective nature of the study.

Sources of funding

This study was supported by the National Key R&D Program of China (Program No. 2020YFC2009002) and National Natural Science Foundation of China (Grant Nos. 82371208 and 81974160).

Author contribution

R.S.: conceptualization, data curation, formal analysis, software, validation, visualization, and writing – original draft; S.L. and Y. W.: formal analysis, software, validation, visualization, and writing – review and editing; L.H.: methodology, visualization, and writing – review and editing; Q.X.: investigation, visualization, and writing – review and editing; G.Z., Y.H., and X.Y.: investigation and writing – review and editing; Y.W.: formal analysis, software, and methodology; X.L.: conceptualization, resources, supervision, and writing – review and editing; A.L.: funding acquisition, project administration, supervision, and writing – review and editing; Z.Z.: conceptualization, data curation, resources, methodology, project administration, and writing – review and editing.

Conflicts of interest disclosure

The authors declare that they have no conflicts of interest.

Research registration unique identifying number (UIN)

1. Name of the registry: Clinicaltrials.gov.
2. Unique identifying number or registration ID: NCT06146829.
3. 3.Hyperlink to your specific registration (must be publicly accessible and will be checked): https://clinicaltrials.gov/ct2/show/NCT06146829.

Guarantor

Zhiqiang Zhou and Ailin Luo had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Data availability statement

Data underlying this article will be uploaded to https://data center.tjh.com.cn, and will be made available on reasonable request to the corresponding author (Zhiqiang Zhou).

Provenance and peer review

Not commissioned, externally peer-reviewed.

References

[1] Pang Z, Liang S, Xing M, *et al*. The correlation of intraoperative oliguria with acute kidney injury after noncardiac surgery: a systematic review and meta-analysis. Int J Surg 2023;109:449–57.
[2] Sun R, Zhou Z, Li X, *et al*. Prognostic significance of preoperative nutritional status for postoperative acute kidney injury in older patients undergoing major abdominal surgery: a retrospective cohort study. Int J Surg 2024;110:873–83.
[3] French WB, Shah PR, Fatani YI, *et al*. Mortality and costs associated with acute kidney injury following major elective, non-cardiac surgery. J Clin Anesth 2022;82:110933.
[4] Mok V, Nixon J, Hu J, *et al*. The impact of perioperative acute kidney injury/failure on short and long surgical outcomes. Anesthesiol Perioperat Sci 2023;1:9.
[5] Hoste EA, Bagshaw SM, Bellomo R, *et al*. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. Intensive Care Med 2015;41:1411–23.
[6] Shamout F, Zhu T, Clifton DA. Machine learning for clinical outcome prediction. IEEE Rev Biomed Eng 2021;14:116–26.
[7] Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol 2019;20:e262–73.
[8] Nohara Y, Matsumoto K, Soejima H, *et al*. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. Comput Methods Programs Biomed 2022;214:106584.
[9] Ren Y, Loftus TJ, Datta S, *et al*. Performance of a machine learning algorithm using electronic health record data to predict postoperative complications and report on a mobile platform. JAMA Netw Open 2022; 5:e2211973.
[10] 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:e212240.
[11] Hofer IS, Lee C, Gabel E, *et al*. Development and validation of a deep neural network model to predict postoperative mortality, acute kidney injury, and reintubation using a single feature set. NPJ Digit Med 2020;3:58.
[12] Satapathy P, Pradhan KB, Rustagi S, *et al*. Application of machine learning in surgery research: current uses and future directions - editorial. Int J Surg 2023, 109:1550–1.
[13] Fransvea P, Fransvea G, Liuzzi P, *et al*. Study and validation of an explainable machine learning-based mortality prediction following emergency surgery in the elderly: a prospective observational study. Int J Surg 2022;107:106954.
[14] Lei VJ, Luong T, Shan E, *et al*. Risk stratification for postoperative acute kidney injury in major noncardiac surgery using preoperative and intraoperative data. JAMA Netw Open 2019;2:e1916921.
[15] 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;24:478.
[16] Chen Z, Chen L, Yao G, *et al*. Novel blood cytokine-based model for predicting severe acute kidney injury and poor outcomes after cardiac surgery. J Am Heart Assoc 2020;9:e018004.
[17] 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;7:322.
[18] Thongprayoon C, Pattharanitima P, Kattah AG, *et al*. Explainable preoperative automated machine learning prediction model for cardiac surgery-associated acute kidney injury. J Clin Med 2022;11:6264.
[19] Nikkinen O, Kolehmainen T, Aaltonen T, *et al*. Developing a supervised machine learning model for predicting perioperative acute kidney injury in arthroplasty patients. Comput Biol Med 2022;144:105351.
[20] Lee HC, Yoon S, Yang SM, *et al*. Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. J Clin Med 2018;7:428.
[21] Wu M, Jiang X, Du K, *et al*. Ensemble machine learning algorithm for predicting acute kidney injury in patients admitted to the neurointensive care unit following brain surgery. Sci Rep 2023;13:6705.
[22] Mathew G, Agha R, Albrecht J, *et al*. STROCSS 2021: strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. Int J Surg 2021;96:106165.
[23] Lameire NH, Levin A, Kellum JA, *et al*. Harmonizing acute and chronic kidney disease definition and classification: report of a Kidney Disease: Improving Global Outcomes (KDIGO) Consensus Conference. Kidney Int 2021;100:516–26.
[24] Alkharusi H. Categorical variables in regression analysis: a comparison of dummy and effect coding. Int J Educat 2012;4:202–10.
[25] Bihorac A, Ozrazgat-Baslanti T, Ebadi A, *et al*. MySurgeryRisk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg 2019;269:652–62.
[26] McKinlay J, Tyson E, Forni LG. Renal complications of anaesthesia. Anaesthesia 2018;73(suppl 1):85–94.
[27] Prowle JR, Forni LG, Bell M, *et al*. Postoperative acute kidney injury in adult non-cardiac surgery: joint consensus report of the Acute Disease

Quality Initiative and PeriOperative Quality Initiative. Nat Rev Nephrol 2021;17:605–18.
[28] Vaara ST, Bellomo R. Postoperative renal dysfunction after noncardiac surgery. Curr Opin Crit Care 2017;23:440–6.
[29] Iyigun M, Aykut G, Tosun M, *et al*. Perioperative risk factors of acute kidney injury after non-cardiac surgery: a multicenter, prospective, observational study in patients with low grade American Society of Anesthesiologists physical status. Am J Surg 2019;218:457–61.
[30] Grams ME, Sang Y, Coresh J, *et al*. Acute kidney injury after major surgery: a retrospective analysis of veterans health administration data. Am J Kidney Dis 2016;67:872–80.
[31] Sun K, Li F, Lin D, *et al*. Serum gamma - glutamyltransferase is associated with albuminuria: a population-based study. PLoS One 2014;9:e114970.
[32] Yilmaz MI, Turgut F, Kanbay M, *et al*. Serum gamma-glutamyltransferase levels are inversely related to endothelial function in chronic kidney disease. Int Urol Nephrol 2013;45:1071–8.
[33] Javid M, Mirdamadi A, Javid M, *et al*. Gamma glutamyl transferase as a biomarker to predict contrast-induced nephropathy among patients with acute coronary syndrome undergoing coronary interventions: a meta-analysis. Ann Med Surg (Lond) 2023;85:4033–40.
[34] Limaye K, Yang JD, Hinduja A. Role of admission serum albumin levels in patients with intracerebral hemorrhage. Acta Neurol Belg 2016;116:27–30.
[35] Li N, Qiao H, Guo JF, *et al*. Preoperative hypoalbuminemia was associated with acute kidney injury in high-risk patients following non-cardiac surgery: a retrospective cohort study. BMC Anesthesiol 2019;19:171.
[36] Lee EH, Baek SH, Chin JH, *et al*. Preoperative hypoalbuminemia is a major risk factor for acute kidney injury following off-pump coronary artery bypass surgery. Intensive Care Med 2012;38:1478–86.
[37] Kertai MD, Zhou S, Karhausen JA, *et al*. Platelet counts, acute kidney injury, and mortality after coronary artery bypass grafting surgery. Anesthesiology 2016;124:339–52.
[38] Aghdaii N, Ferasatkish R, Mohammadzadeh Jouryabi A, *et al*. Significance of preoperative total lymphocyte count as a prognostic criterion in adult cardiac surgery. Anesth Pain Med 2014;4:e20331.
[39] Nangaku M. Chronic hypoxia and tubulointerstitial injury: a final common pathway to end-stage renal failure. J Am Soc Nephrol 2006;17:17–25.
[40] Wu Y, Chen J, Luo C, *et al*. Predicting the risk of postoperative acute kidney injury: development and assessment of a novel predictive nomogram. J Int Med Res 2021;49:3000605211032838.
[41] De Rosa S, Greco M, Rauseo M, *et al*. The good, the bad, and the serum creatinine: exploring the effect of muscle mass and nutrition. Blood Purif 2023;52:775–85.
[42] Bang JY, Jun IG, Lee JB, *et al*. Impact of sarcopenia on acute kidney injury after infrarenal abdominal aortic aneurysm surgery: a propensity matching analysis. Nutrients 2021;13:2212.
[43] Prowle JR, Croal B, Abbott TEF, *et al*. investigators Ms. Cystatin C or creatinine for pre-operative assessment of kidney function and risk of post-operative acute kidney injury: a secondary analysis of the METS cohort study. Clin Kidney J 2024;17:sfae004.
[44] Kanda E, Muneyuki T, Kanno Y, *et al*. Uric acid level has a U-shaped association with loss of kidney function in healthy people: a prospective cohort study. PLoS One 2015;10:e0118031.
[45] Park JH, Jo YI, Lee JH. Renal effects of uric acid: hyperuricemia and hypouricemia. Korean J Intern Med 2020;35:1291–304.
[46] Mendez CE, Der Mesropian PJ, Mathew RO, *et al*. Hyperglycemia and acute kidney injury during the perioperative period. Curr Diab Rep 2016;16:10.
[47] Jiang J, Li S, Zhao Y, *et al*. Intensive glucose control during the perioperative period for diabetic patients undergoing surgery: an updated systematic review and meta-analysis. J Clin Anesth 2021;75:110504.
[48] Lee YL, Chen BK, Lin KD, *et al*. The impact of severe hypoglycemia on renal impairment in type 2 diabetes. Diabetes Res Clin Pract 2015;108:448–55.
[49] Yun JS, Park YM, Han K, *et al*. Severe hypoglycemia and the risk of end stage renal disease in type 2 diabetes. Sci Rep 2021;11:4305.
[50] Boon AC, Bulmer AC, Coombes JS, *et al*. Circulating bilirubin and defense against kidney disease and cardiovascular mortality: mechanisms contributing to protection in clinical investigations. Am J Physiol Renal Physiol 2014;307:F123–36.
[51] Kumar A, Pant P, Basu S, *et al*. Oxidative stress in neonatal hyperbilirubinemia. J Trop Pediatr 2007;53:69–71.
[52] Wu YH, Wu CY, Cheng CY, *et al*. Severe hyperbilirubinemia is associated with higher risk of contrast-related acute kidney injury following contrast-enhanced computed tomography. PLoS One 2020;15:e0231264.
[53] Zarbock A, Weiss R, Albert F, *et al*. Epidemiology of surgery associated acute kidney injury (EPIS-AKI): a prospective international observational multi-center clinical study. Intensive Care Med 2023;49:1441–55.
[54] Chiu C, Fong N, Lazzareschi D, *et al*. Fluids, vasopressors, and acute kidney injury after major abdominal surgery between 2015 and 2019: a multicentre retrospective analysis. Br J Anaesth 2022;129:317–26.
[55] Karkouti K, Grocott HP, Hall R, *et al*. Interrelationship of preoperative anemia, intraoperative anemia, and red blood cell transfusion as potentially modifiable risk factors for acute kidney injury in cardiac surgery: a historical multicentre cohort study. Can J Anaesth 2015;62:377–84.
[56] Zhu S, Lu P, Liu Z, *et al*. Longitudinal hemoglobin trajectories and acute kidney injury in patients undergoing cardiac surgery: a retrospective cohort study. Front Cardiovasc Med 2023;10:1181617.
[57] Zhou Y, Liu S. Threshold heterogeneity of perioperative hemoglobin drop for acute kidney injury after noncardiac surgery: a propensity score weighting analysis. BMC Nephrol 2022;23:206.