Building a Prediction Model for Postoperative Acute Kidney Injury using Machine Learning: The CMC-AKIX Model

Hye Eun Yoon, Jiwon Min, Jae-Hong Min, Se-Hyun Chang, Byung Ha Chung, Eun Sil Koh, Young Soo Kim, Hyung Wook Kim, Tae Hyun Ban, Seok Joon Shin, In Young Choi

Submitted to: Journal of Medical Internet Research
on: June 10, 2024

Table of Contents

Building a Prediction Model for Postoperative Acute Kidney Injury using Machine Learning: The CMC-AKIX Model

Hye Eun Yoon[1*] MD, PhD; Jiwon Min[2*] MD, PhD; Jae-Hong Min[3*] MD; Se-Hyun Chang[4]; Byung Ha Chung[5] MD, PhD; Eun Sil Koh[6] MD, PhD; Young Soo Kim[7] MD, PhD; Hyung Wook Kim[8] MD, PhD; Tae Hyun Ban[9] MD, PhD; Seok Joon Shin[1] MD, PhD; In Young Choi[10*] PhD

[1]Department of Internal Medicine, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea Incheon KR
[2]Department of Internal Medicine, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea Seoul KR
[3]School of Information, University of California Berkley US
[4]Department of Medical Informatics, College of Medicine The Catholic University of Korea Seoul KR
[5]Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea Seoul KR
[6]Department of Internal Medicine, Yeouido St. Mary's Hospital College of Medicine, The Catholic University of Korea Seoul KR
[7]Department of Internal Medicine, Uijeongbu St. Mary's Hospital, College of Medicine The Catholic University of Korea Seoul KR
[8]Department of Internal Medicine, St. Vincent's Hospital, College of Medicine The Catholic University of Korea Seoul KR
[9]8Department of Internal Medicine, Eunpyeong St. Mary's Hospital, College of Medicine The Catholic University of Korea Seoul KR
[*]these authors contributed equally

Corresponding Author:
In Young Choi PhD

Abstract

Background: Postoperative acute kidney injury (AKI) is a significant risk associated with surgeries under general anesthesia, often leading to increased mortality and morbidity. Existing predictive modelWe proposed to build a prediction model for postoperative AKI using several machine learning methods.s for postoperative AKI are usually limited to specific surgical areas or require external validation.

Objective: We proposed to build a prediction model for postoperative AKI using several machine learning methods.

Methods: We conducted a retrospective cohort analysis of noncardiac surgeries from 2009 to 2019 at seven university hospitals in South Korea. We evaluated six machine learning models: deep neural networks, logistic regression, decision tree, random forest, light gradient boosting machine, and naïve Bayes for predicting postoperative AKI, defined as a significant increase in serum creatinine or initiation of renal replacement therapy within 30 days post-surgery.

Results: Among 239,267 surgeries analyzed, 7,935 cases of postoperative AKI were identified. Performance of the models were analyzed using area under the curve (AUC) of receiver operating characteristic (ROC) curve, accuracy, precision, sensitivity (recall), specificity and F1 score. The models, utilizing 38 preoperative predictors, showed that deep neural networks (AUC = 0.832), light gradient boosting machines (AUC = 0.836), and logistic regression (AUC = 0.825) demonstrated superior performance in predicting AKI risk. The model with the best performance was then developed into a user-friendly website for clinical use.

Conclusions: Our study introduces a robust, high-performance AKI risk prediction system applicable in clinical settings using preoperative data. This model's integration into a user-friendly website enhances its clinical utility, offering a significant step forward in personalized patient care and risk management.

(JMIR Preprints 10/06/2024:62853)
DOI: https://doi.org/10.2196/preprints.62853

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
- ✔ **Please make my preprint PDF available to anyone at any time (recommended).**
- Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
- Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Building a Prediction Model for Postoperative Acute Kidney Injury using Machine Learning: The CMC-AKIX Model

Running title: Prediction Model for Postoperative Acute Kidney Injury

Jiwon Min MD, PhD[1*], Jae-Hong Min MD[2*], Se-Hyun Chang[3], Byung Ha Chung MD, PhD[4], Eun Sil Koh MD, PhD[5], Young Soo Kim MD, PhD.[6], Hyung Wook Kim MD, PhD[7], Tae Hyun Ban MD, PhD.[8], Seok Joon Shin MD, PhD.[9], In Young Choi PhD.[3†], Hye Eun Yoon MD, PhD[9†]

[1]Department of Internal Medicine, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[2]School of Information, University of California, Berkley, USA
[3]Department of Medical Informatics, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[4]Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[5]Department of Internal Medicine, Yeouido St. Mary's Hospital College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[6]Department of Internal Medicine, Uijeongbu St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[7]Department of Internal Medicine, St. Vincent's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[8]Department of Internal Medicine, Eunpyeong St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[9]Department of Internal Medicine, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

[*] Jiwon Min and Jae-Hong Min contributed equally to this work as first authors.
[†] In Young Choi and Hye Eun Yoon contributed equally to this work as corresponding authors.

[*]Correspondence:
Hye Eun Yoon, MD, PhD
Division of Nephrology, Department of Internal Medicine, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea
56, Dongsu-ro, Bupyeong-gu, Incheon, 21431, Republic of Korea.
Telephone: +82 32 280 7370; Fax: +82 32 280 5987; E-mail: berrynana@catholic.ac.kr

In Young Choi, PhD
Department of Medical Informatics, Graduate School of Healthcare Management & Policy, College of Medicine, The Catholic University of Korea
Tel: 82-2-2258-7870
E-mail: iychoi@catholic.ac.kr

Abstract

Background: Postoperative acute kidney injury (AKI) is a significant risk associated with surgeries under general anesthesia, often leading to increased mortality and morbidity. Existing predictive models for postoperative AKI are usually limited to specific surgical areas or require external validation.

Objective: We proposed to build a prediction model for postoperative AKI using several machine learning methods.

Methods: We conducted a retrospective cohort analysis of noncardiac surgeries from 2009 to 2019 at seven university hospitals in South Korea. We evaluated six machine learning models: deep neural networks, logistic regression, decision tree, random forest, light gradient boosting machine, and naïve Bayes for predicting postoperative AKI, defined as a significant increase in serum creatinine or initiation of renal replacement therapy within 30 days post-surgery.

Results: Among 239,267 surgeries analyzed, 7,935 cases of postoperative AKI were identified. Performance of the models were analyzed using area under the curve (AUC) of receiver operating characteristic (ROC) curve, accuracy, precision, sensitivity (recall), specificity and F1 score. The models, utilizing 38 preoperative predictors, showed that deep neural networks (AUC = 0.832), light gradient boosting machines (AUC = 0.836), and logistic regression (AUC = 0.825) demonstrated superior performance in predicting AKI risk. The model with the best performance was then developed into a user-friendly website for clinical use.

Conclusions: Our study introduces a robust, high-performance AKI risk prediction system applicable in clinical settings using preoperative data. This model's integration into a user-friendly website enhances its clinical utility, offering a significant step forward in personalized patient care and risk management.

KEY WORDS

acute kidney injury; general surgery; deep neural networks; machine learning; prediction model

Introduction

Acute kidney injury (AKI) represents a critical challenge in postoperative care, significantly affecting patient outcomes and healthcare systems. It is a common complication that affects up to 5-7.5% of total hospitalized patients, with a markedly higher prevalence of 20% in intensive care units (ICUs) [1]. Among AKI of hospitalized patients, postoperative AKI impacts up to 40% of patients following surgery [1]. This condition not only escalates morbidity but also substantially increases in-hospital mortality to approximately 3 to 9-fold [2]. The severity of this risk is further underscored in patients who developed postoperative AKI after intraabdominal surgery, as a large-scale study reported a 15-fold higher risk of mortality in patients with AKI compared to those without AKI [3]. Moreover, even patients whose renal function completely recovered after postoperative AKI still faced higher risk of death compared to those without AKI [4], highlighting the profound and lasting consequences of this condition. These statistics underscore the need for accurate prediction and preemptive management of AKI in the postoperative setting.

There are many factors associated with postoperative AKI; age, sex, obesity, type of surgery, medications including renin-angiotensin-aldosterone system inhibitors (RASi) and non-steroidal anti-inflammatory drugs (NSAIDs), and comorbidities such as chronic kidney disease (CKD), diabetes, hypertension, cardiovascular disease, liver disease, and chronic obstructive pulmonary disease (COPD) [5-7]. These factors need to be integrated to assess the risk of postoperative AKI before surgery, and accurate risk prediction enables recognition of patients who need pre-, intra-, and postoperative management to alleviate the risk. Several risk scoring tools for postoperative AKI have been described [8-10]. However, their limitations are heterogeneity of the study population, inclusion of a single center or a small number of centers, and lack of external validation. To make the risk scoring system generalizable, a validation from a larger cohort using a multicenter database is needed [11]. Machine learning allows greater insight into possible interactions between variables and searches for as many informative and interesting feature relationships, including those in subgroups, which can discover new variables involved in the event and is useful in a large dataset [12]. Therefore, the aim of this study was to build a risk prediction model for postoperative AKI using machine learning methods from a multicenter cohort.

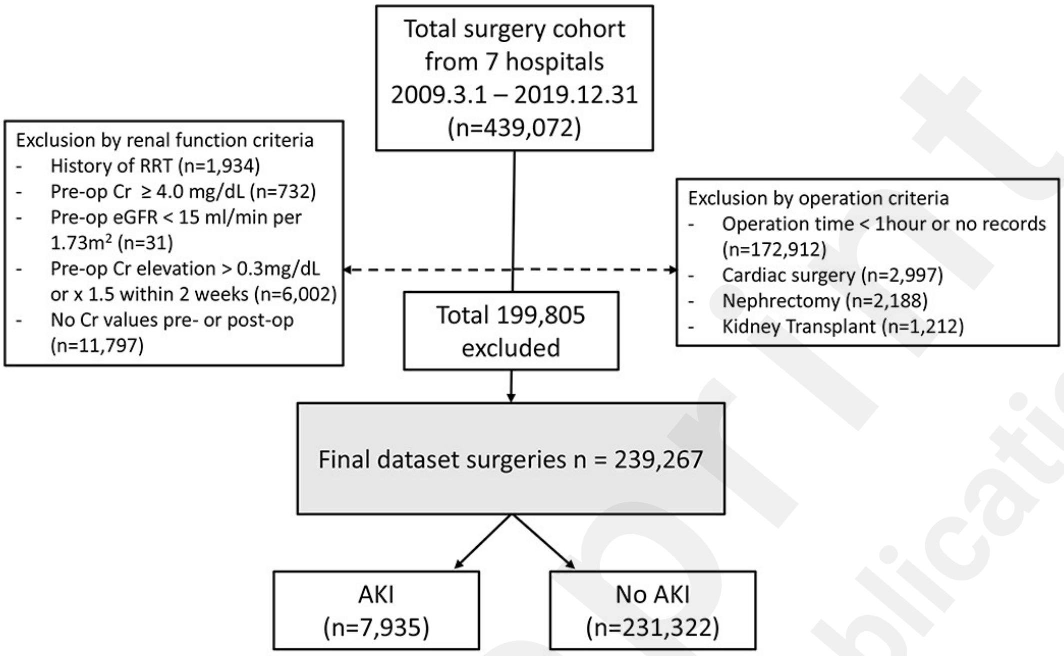
Methods

Study Population

Patients who underwent general anesthesia surgery from 1st March 2009 to 31st December 2019 at 7 academic hospitals of the Catholic University of Korea (Seoul St. Mary's, Yeouido St. Mary's, Uijeongbu St. Mary's, Eunpyeong St. Mary's, Bucheon St. Mary's, St Vincent, and Incheon St. Mary's Hospitals) were included. The exclusion criteria were as follows: operation-related criteria were operation duration under 1 hour or duration not available, cardiac surgeries, operations of brain death donors, nephrectomies, kidney transplant operations. Renal function-related exclusion criteria were patients with history of renal replacement therapy (RRT), preoperative serum creatinine (sCr) ≥ 4.0 mg/dL or estimated glomerular filtration rate (eGFR) <15 ml/min per 1.73 m^2, elevation of preoperative sCr more than 0.3 mg/dL or 1.5 times within 2 weeks before surgery, and patients without preoperative or postoperative sCr values (Figure 1). The study was approved from the institutional review board of the Catholic University of Korea, College of Medicine

(XC20WIDI0080) with waiver of consent due to the retrospective study methods. This report has been written according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement [13].

Figure 1. Flow Chart of the Study Population. AKI: acute kidney injury; CMC: Catholic Medical Center; Cr: creatinine; eGFR: estimated glomerular filtration rate; RRT, renal replacement therapy.



Definition of Postoperative AKI

Postoperative AKI was defined as AKI that developed within 30 days after surgery, using the Kidney Disease: Improving Global Outcomes (KDIGO) criteria [14]. Stage 1 AKI was defined as sCr 1.5 to 1.9 times above baseline or an increase in sCr ≥ 0.3 mg/dL, stage 2 AKI was defined as sCr 2.0 to 2.9 times above baseline, and stage 3 AKI was defined as sCr more than 3 times above baseline or ≥ 4 mg/dL or initiation of renal replacement therapy (hemodialysis, peritoneal dialysis, or continuous renal replacement therapy). We did not use the urine output criteria of KDIGO as previous studies suggested that the threshold of oliguria for postoperative AKI may be different from those of other AKIs [15,16] and lack of data. This definition of postoperative AKI was used to create the supervised learning dataset of those with or without postoperative AKI.

Datal Collection and Cleansing

We collected data on demographic characteristics, underlying clinical diseases, preoperative laboratory data, preoperative medication, and surgical characteristics such as expected operation time, the day of operation (weekday or weekend), and department of surgery. Underlying diseases of subjects were determined using *International Classification of Disease, Tenth Revision* (*ICD-10*) codes of principal and secondary diagnosis. Comorbid diseases and ICD-10 codes are shown in Table S1 in Multimedia Appendix 1. Preoperative medications included RASi [angiotensin converting enzyme inhibitor (ACEi) or angiotensin II type 1 receptor blocker (ARB)] or NSAIDs. Preoperative eGFR was calculated from the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [17]. Body mass index (BMI) was calculated as the patient's weight in kilograms divided by height in meters squared (kg/m^2).

Data was extracted from the Catholic Medical Center-Clinical Data Warehouse (CMC-CDW) and processed using R software version 3.6.3. The 38 variables included in the final analysis are shown in Table 1. Data artifacts and extreme values were set to the first percentile and 99th percentile and missing values were filled using Multiple Imputation by Chained Equations (MICE) [18]. MICE was used to provide more accurate estimates of the missing variables with the correlation of missing variables to other existing data points [19]. Non-binary data was one hot encoded (OHE), a method for rearranging categorical data into binary variables, and numerical data was normalized using Min-Max Scaling. This would convert all numeric values between or equal to a value of 0 and 1. Mix-Max Scaling is given by:

$$X_{\text{Scaled}} = \frac{X - X_{\text{Min}}}{X_{\text{Max}} - X_{\text{Min}}}$$

OHE, Min-Max Scaling and dataset splitting was accomplished using Scikit-Learn Library version 0.24.2 [20]. These steps are required to improve the performance of machine learning models and training stability. During feature selection, the hematocrit variable was removed because it had > 0.9 correlation with preoperative hemoglobin levels. Because there was a small percentage of AKI events (3.3%), there was an extreme class imbalance in the dataset. Such imbalances can cause a falsely elevated accuracy and adversely affect machine learning training [21]. To help overcome this issue, the AKI training dataset was augmented using an oversampling method by synthetic minority over-sampling technique (SMOTE) which has shown to improve imbalanced class classifications (using imblearn library version 0.8.0) [22,23]. The dataset was divided into the training set (80%) and test set (20%). The training set (n = 191,413) and test set (n = 47,854) were balanced for outcomes and randomly assigned.

Table 1. Variables included in the final analysis.

Patient parameters	Surgical parameters	Laboratory parameters
Age	Department	White blood cell count
Sex	Weekday	Hemoglobin
Systolic BP	Operation duration	C-reactive protein
Diastolic BP		Glucose
BMI		Urea nitrogen
Chronic kidney disease		Creatinine
Diabetes		eGFR
Hypertension		Total Protein
Cerebrovascular disease		Albumin
Coronary artery disease		AST
COPD		ALT
Liver cirrhosis		Sodium
Smoking		Potassium
Preoperative ACEi or ARB usage		Chloride
Preoperative NSAID usage		Calcium
		Uric Acid
		Creatine phosphokinase
		Lactic dehydrogenase
		Urine specific gravity
		Urine protein

BP: blood pressure; BMI: body mass index; COPD: chronic obstructive pulmonary disease; ACEi: angiotensin converting enzyme inhibitor; ARB: angiotensin II type 1 receptor blocker; NSAID: non-steroidal anti-inflammatory drug; eGFR: estimated glomerular filtration rate; AST: aspartate aminotransferase; ALT: alanine aminotransferase.

Machine Learning

Various machine learning methods were used to create the model which was trained and evaluated using Python version 3.8.5. Machine learning methods commonly used in healthcare were applied [24,25]. Models applied were logistic regression, decision tree, random forest, naïve Bayes (using Scikit-Learn Library version 0.24.2) [20], light gradient boosting machine (GBM) (using lightgbm version 3.2.1) [26] and deep neural networks (DNN) (using Keras Library version 2.5.0) [27]. The strengths and weaknesses of each model have been summarized in Table 2.

For the deep learning model, the structure that was chosen was a model that had an input layer of width 50 (to account for the 40 inputs and to include the One Hot Encodings), 3 hidden layers with a width of 64, 32 and 32 and a single output node. The configurations of the different models are shown Table S2 in Multimedia Appendix 2. The training of the models was done on a machine with Intel Xeon Gold 6240R (8 cores) at 2.40 GHz, with 64GB of RAM, Windows 10 Enterprise Build 17763. To analyze the statistical performance of the models for postoperative AKI prediction, we assessed area under the curve (AUC) of receiver operating characteristic (ROC) curve, accuracy, precision, sensitivity (recall), specificity and F1 score.

Table 2. Characteristics of machine learning methods.

Method	How it works	Advantages	Disadvantages
DNN	Multiple layers of interconnected nodes (neurons) of at least 3 hidden layers or more. Each neuron is a weighted sum of inputs and produces output by an activation function. Learns by backpropagation.	Can capture complex relationships between features, especially in larger datasets. Can capture hierarchical features.	Requires large amounts of data to avoid overfitting. Computationally expensive.
Logistic Regression	Linear classification algorithm that finds relationships between independent variables and a binary outcome using the probability from logistical functions.	Computationally less intensive. Large datasets can be reasonably adapted.	May not capture complex relationships between features.
Decision Tree	A number of nodes which separate features depending on feature values and continues at each node, representing a tree.	Can capture complex relationships between features, especially in larger datasets.	Prone to overfitting of data. Can produce biased decision tree depending on features.
Random Forest	An ensemble (group) of decision trees that randomly select features and data for training, with decisions made by the ensemble using regression or other methods.	Can capture complex relationships between features, especially in larger datasets. Less likely to overfit compared to a single decision tree.	Can be computationally expensive.
Light GBM	An ensemble (group) of "weak models" (usually decision trees) which are sequentially added to one another to help to improve performance over a number of iterations.	Can capture complex relationships between features, especially in larger datasets. Can capture more complex relationships compared to Random Forest.	Can possibly overfit data.
Naïve Bayes	Makes use of conditional probability to represent the	Computationally less intensive.	May not capture complex

	likelihood of classification given a certain set of features, assuming that each feature is independent of one another.	Large datasets can be reasonably adapted.	relationships between features. Relies on the assumption that features are independent of one another.

DNN: deep neural network; GBM: gradient boosting machine.

Statistical Analysis

Statistical analysis was performed using SAS software, version 9.4 (SAS Institute Inc. Cary, NC, USA). Continuous variables were presented as mean and SD for data with normal distribution and presented as median and interquartile ranges for data with nonparametric distribution. After distribution of data between two groups was determined, they were compared using independent t-test or Wilcoxon’s rank sum test. Categorical data was presented as percentages and comparison between the two groups was performed using Chi-square test, or Fisher’s exact test. To determine the risk factors for AKI we used the logistic regression model. Multivariable analysis using logistic regression was performed on variables with a *P*-value < .2 on univariable analysis. The results are presented as odds ratio (OR) with 95% CI. *P*-values < .05 were considered significant.

Results

Patient baseline characteristics

A total of 439,072 surgery cases from 7 academic hospitals of the Catholic University of Korea were included in the study (Figure 1). After exclusion of patients according to the exclusion criteria mentioned above, a total of 239,267 datasets were included in the final analysis. Among which 7,935 (3.3%) AKI events occurred. Baseline demographics of AKI versus No AKI patients are shown in Table 3. Significant differences were observed in all baseline characteristics between the two groups. AKI patients were older, with higher percentage of those of the female sex, and had lower BMI, higher percentage of smokers, and higher baseline systolic and diastolic blood pressure. The AKI group also showed higher percentage of all preexisting comorbidities (CKD, diabetes, hypertension, coronary artery disease, cerebrovascular disease, COPD, or liver cirrhosis), more frequent usage of RASi (ACEi or ARB) and NSAIDs, and higher percentage of patients undergoing general surgery, neurosurgery, and thoracic surgery. Laboratory results of the AKI group showed lower levels of hemoglobin and serum albumin and eGFR, and higher baseline sCr and C-reactive protein levels.

Table 3. Baseline characteristics

Variables	No AKI (n = 231,322)	AKI (n = 7,935)	*P*-value
Age, years	54.56 ± 15.16	61.36 ± 14.19	< .001
Male sex, n (%)	126,217 (54.6)	2,997 (37.8)	< .001
BMI, kg/m^2	24.29 ± 4.06	24.07 ± 4.22	< .001
Smoker, n (%)	34,228 (14.8)	1,424 (18.0)	< .001
Preexisting comorbidities, n (%)			
Chronic kidney disease	298 (0.1)	114 (1.4)	< .001
Diabetes	7,295 (3.2)	2663 (8.4)	< .001
Hypertension	5,802 (2.5)	504 (6.4)	< .001
Coronary artery disease	3,927 (1.7)	308 (3.9)	< .001
Cerebrovascular disease	11,960 (5.2)	734 (9.3)	< .001
COPD	952 (0.4)	85 (1.1)	< .001
Liver Cirrhosis	1,077 (0.5)	131 (1.9)	< .001
Systolic BP, mmHg	125.84 ± 15.74	133.81 ± 20.70	< .001
Diastolic BP, mmHg	69.71 ± 9.99	67.01 ± 12.69	< .001
Medication, n (%)			
ACEi or ARB	7,844 (3.4)	768 (9.7)	< .001
NSAIDs	35,006 (15.1)	1,650 (20.8)	< .001
Department, n (%)			< .001
General Surgery	61,948 (26.8)	2,409 (30.4)	
Neurosurgery	31,169 (13.5)	1,407 (17.7)	
Orthopedics	60,442 (26.1)	1176 (14.8)	
Obstetrics and gynecology	22,607 (9.8)	299 (3.8)	
Otorhinolaryngology	11,551 (5.0)	148 (1.9)	
Thoracic surgery	10,774 (4.7)	427 (5.4)	
Others	32,831 (14.2)	2,069 (26.1)	
Preoperative laboratory results			
Hemoglobin, g/dL	13.14 ± 1.83	12.10 ± 2.15	< .001
Urea nitrogen, mg/dL	14.69 ± 5.47	16.86 ± 8.75	< .001
Creatinine, mg/dL	0.81 ± 0.24	0.91 ± 0.43	< .001
eGFR, mL/min/1.73m^2	95.41 ± 23.22	82.75 ± 27.12	< .001
Albumin, g/dL	4.16 ± 0.51	3.67 ± 0.69	< .001
C-reactive protein, g/dL	4.04 ± 18.35	8.40 ± 26.40	< .001

BMI: body mass index; COPD: chronic obstructive pulmonary disease; BP: blood pressure; ACEi: angiotensin converting enzyme inhibitor; ARB: angiotensin II type 1 receptor blocker;

NSAID: non-steroidal anti-inflammatory drug; eGFR: estimated glomerular filtration rate.

Model Performance

Variable selection was performed using logistic regression (clinical parameters are shown in Table S3 in Multimedia Appendix 3 and laboratory parameters are shown in Table S4 in Multimedia Appendix 4). The performances of the different models are shown in Table 4. We hypothesized that a simple system not using too many variables, e.g. less than 20 variables, would be more practical to use in a clinical setting. Therefore, we evaluated Model 2 and Model 3 using multiple machine learning methods. Model 2 included 11 variables which was used in the classification system developed by Park et al [8], including age, sex, emergency operation, operation duration, diabetes, ACEi or ARB usage, blood levels of albumin, hemoglobin and sodium, eGFR and urine dipstick protein. In this model DNN showed the highest performance (AUC=0.800). Model 3 included variables that were found significant on multivariable analysis, including age, sex, systolic blood pressure, diastolic blood pressure, operation duration, eGFR, blood levels of creatinine, albumin, sodium, potassium, chloride, glucose, and lactic dehydrogenase, and urine dipstick protein. In this model as well, DNN showed the highest performance (AUC=0.811). Model 1 included all 38 preoperative variables and surgical characteristics, on which light GBM (AUC=0.836), DNN (AUC=0.832) and logistic regression (AUC=0.825) demonstrated the best prediction performance (shown in Figure 2).

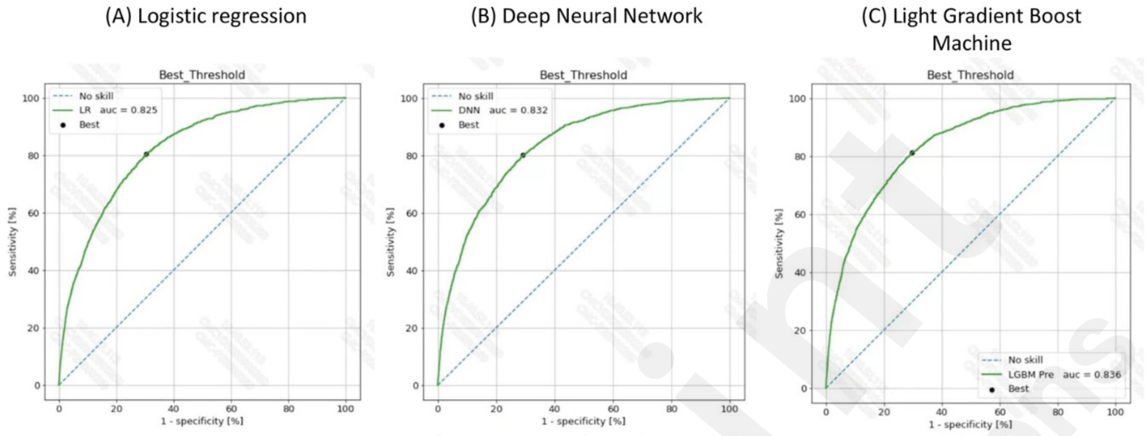
Table 4. Performance metrics of postoperative AKI prediction models.

Analysis	Model	AUC	Accuracy	Precision	Specificity	Recall/ Sensitivity	F1 score
DNN	Model 1	0.832	0.711	0.086	0.708	0.802	0.156
	Model 2	0.800	0.712	0.082	0.711	0.750	0.147
	Model 3	0.811	0.691	0.079	0.688	0.785	0.144
Logistic Regression	Model 1	0.825	0.700	0.083	0.696	0.805	0.151
	Model 2	0.790	0.730	0.083	0.731	0.709	0.148
	Model 3	0.806	0.719	0.083	0.718	0.741	0.149
Decision Tree	Model 1	0.679	0.606	0.056	0.603	0.691	0.104
	Model 2	0.711	0.635	0.062	0.633	0.708	0.114
	Model 3	0.626	0.844	0.085	0.860	0.379	0.139
Random Forest	Model 1	0.813	0.751	0.092	0.752	0.732	0.163
	Model 2	0.806	0.708	0.081	0.706	0.751	0.146
	Model 3	0.812	0.756	0.091	0.758	0.708	0.161
Light GBM	Model 1	0.836	0.711	0.087	0.708	0.813	0.157
	Model 2	0.810	0.730	0.086	0.730	0.740	0.154
	Model 3	0.825	0.728	0.087	0.727	0.756	0.156
Naïve Bayes	Model 1	0.785	0.680	0.075	0.677	0.767	0.137
	Model 2	0.773	0.662	0.072	0.658	0.770	0.131
	Model 3	0.792	0.742	0.086	0.744	0.701	0.153

AUC: area under the curve; DNN: deep neural network; GBM: gradient boosting machine.
Model 1: age, sex, systolic blood pressure, diastolic blood pressure, body mass index, chronic kidney disease, diabetes mellitus, hypertension, cerebrovascular disease, coronary artery disease, chronic obstructive pulmonary disease, liver cirrhosis, emergency operation, operation duration, ACEi or ARB usage, NSAIDs usage, eGFR, blood levels of creatinine, total protein, albumin, AST, ALT, urea nitrogen, sodium, potassium, chloride, calcium, creatine phosphokinase, lactic dehydrogenase, C-reactive protein, glucose, hemoglobin, and

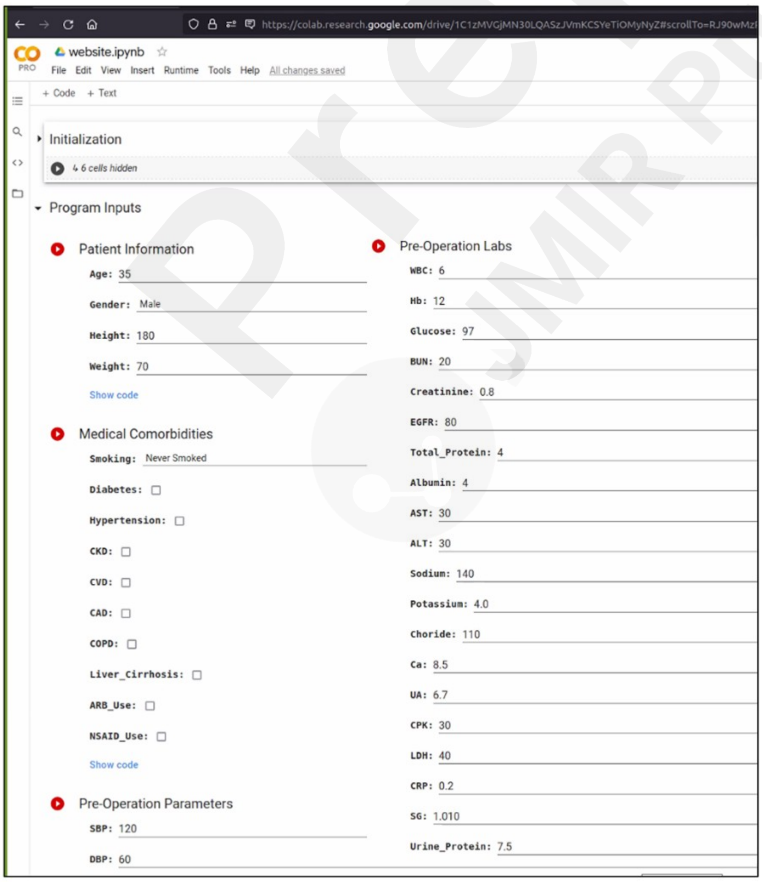
white blood cell count, urine specific gravity and urine protein
Model 2: Age, sex, emergency operation, operation duration, diabetes mellitus, ACEi or ARB usage, blood levels of albumin, hemoglobin, and sodium, eGFR, and urine protein
Model 3: age, sex, systolic blood pressure, diastolic blood pressure, operation duration, eGFR, blood levels of creatinine, albumin, sodium, potassium, chloride, glucose, and lactic dehydrogenase, and urine protein

Figure 2. AUC of best performance models. (A) Light gradient boost machine showed AUC of 0.836, (B) DNN showed AUC of 0.832, and (C) logistic regression showed AUC of 0.825. AUC, area under curve, DNN, deep neural networks; LGBM, light gradient boosting machine.



Therefore, our postoperative AKI prediction tool, the CMC-AKIX, was developed using all 38 variables. Finally, the DNN Model 1 was developed into a user-friendly website which can be accessed at www.cmc-akix.com (shown in Figure 3). This was created using Flask and hosted on a Google Cloud Virtual Machine.

Figure 3. A sample page of the website application



Discussion

Using a multicenter database of 239,267 noncardiac surgeries, we have developed a high-performance risk prediction system for postoperative AKI that can be easily applied. The model uses preoperative patient characteristics and laboratory data along with simple information about the surgery. DNN and light GBM showed a good performance in predicting postoperative AKI, with the best performance when all the 38 variables were included.

AKI has a global presence and a high disease burden and mortality [28]. The incidence of AKI varies widely according to the geographic locations and is dependent on the setting; community-acquired versus hospital-acquired. It was reported that 1 in 5 adults and 1 in 3 children worldwide experience hospital-acquired AKI using the KDIGO definition [29]. Causes of hospital-acquired AKI include sepsis, critical illness, surgery, and use of nephrotoxic medications [30]. Postoperative AKI accounts for 30 – 40% of hospital-acquired AKI [1], and increases the risk of morbidity and in-hospital mortality [2]. Since treatment options are limited, prevention of postoperative AKI is the cornerstone of improving patient outcomes after surgery [1]. Previous studies have found risk factors that increase the risk of postoperative AKI [1,5,6]. However, the definition of AKI using increased sCr levels as a marker of kidney damage has a limitation because sCr levels begin to increase after the pathological changes of kidney injury are already in progress. Therefore, earlier and timely prevention and detection of postoperative AKI can be difficult [31]. This has led to continuous efforts to develop a risk stratification system for postoperative AKI. Recently, Park *et al*. have developed an index to classify postoperative AKI within 90 days after noncardiac surgery from 90,805 patients (SPARK index), which included 11 variables; age, sex, expected surgery duration, emergency operation, diabetes, use of RASi, baseline eGFR, dipstick albuminuria, hypoalbuminemia, anemia, and hyponatremia [8]. The SPARK index showed a discrimination power of AUC 0.80 for postoperative AKI in the discovery cohort and that of AUC 0.72 in the validation cohort.

Machine learning approaches are more flexible than statistical methods as they are free from statistical assumptions such as non-collinearity or normal distribution of residuals. It allows all possible interactions between variables according to multi-dimensional non-linear patterns, and aggressively searches for as many informative and interesting features as possible [12]. Lei *et al*. used machine learning techniques to stratify the risk of postoperative AKI within 7 days after noncardiac surgery from a single center cohort of 42,615 patients [10]. In that study, GBM showed the highest performance with an AUC of 0.817 (95% CI, 0.802 – 0.832), and included 339 preoperative and intraoperative variables. Bihorac *et al.* developed a machine learning-based risk prediction tool (MySurgeryRisk) for 8 major postoperative complications within 24 months after any kind of surgery from a single center cohort of 51,457 patients [9]. Using this platform, the authors validated the model's performance for predicting postoperative AKI, with an AUC of 0.82 (95% CI, 0.82 – 0.83), including 135 variables from a cohort of 22,300 surgeries [32]. The strength of our study is that we used a multi-center dataset of a larger scale than previous ones. Data was extracted from the CMC-CDW, which included data from 7 academic hospitals located in 5 cities in South Korea. The prediction model was developed using 38 clinical and laboratory parameters in combination that exhibited the best prediction performance. These variables are used in clinical practice and can be extracted from electric medical records. In addition, by including the department of surgery as a variable, the CMC-AKIX can be applied to various kinds of noncardiac surgery. We are looking to simplify the model and improve usability by allowing incomplete data or missing values to be filled in with best estimate values using imputational methods such as MICE.

Our study holds a distinct advantage in that it compared several of the most widely employed machine learning methods in clinical data modeling. By doing so, we systematically observed and elucidated the strengths and limitations of each model, using a large, well-curated dataset. We observed that certain methods were more affected by the imbalanced dataset, including the decision tree classifier, random forest, and naïve Bayes. We aim to offer insights in the selection of different algorithms for applications in clinical studies.

This study has several limitations. First, external validation of the results in different countries, race, and ethnicity is needed to make the CMC-AKIX generalizable. Second, our definition of postoperative AKI as AKI developing within 30 days after surgery may be controversial as most studies observing postoperative AKI apply the time period of 7 days in conjunction with the KDIGO criteria [14]. A 30 day-period was selected for this study because postoperative complications or morbidity in most studies is defined as events occurring within 30 days after surgery [33,34]. Patients with AKI that persist for more than 7 days, beyond the 30-day period of the study have been organized in a second cohort study observing postoperative risk of acute kidney disease and CKD [35]. Third, the urine output definition of the KDIGO criteria was not used because of lack of urine output data. This could have led to incomplete identification of postoperative AKI. Fourth, the intraoperative and postoperative factors were not included in the risk prediction system, which also affects postoperative renal outcomes. But since the purpose of our model is mainly to identify patients at high risk for postoperative AKI while they are still in the preoperative setting, intraoperative and postoperative variables should not be included.

In the future, we look to collaborating with other institutions with different demographic data to validate the model and see if it could perform well with different demographic populations. Also, the model will be fine-tuned in the process of including diverse datasets, and performance of the model will be improved by creating an appropriate ensemble of machine learning models to gain the benefits of the different machine learning structures and advantages [36]. Lastly, practical use of the model may be significantly increased by incorporating it into an electronic alert system to automatically identify patients at high risk for postoperative AKI, therefore allowing for proactive management such as cessation of causative medications or prescription of fluids and improving patient care [37].

In conclusion, we propose a machine learning-based risk prediction tool, the CMC-AKIX, using individual patients' preoperative characteristics and surgical information. This model was adapted to a user-friendly online program, and one can use it even if all variables are not included. This tool may guide preoperative counseling and decision making, and perioperative care.

Acknowledgements
This research was supported by Clinical Trials Center of Incheon St. Mary's Hospital, The Catholic University of Korea and by the Institute of Clinical Medicine Research of Bucheon St. Mary's Hospital, Research Fund, 2020. The authors also wish to acknowledge the financial support of the Catholic Medical Center Research Foundation made in the program year of 2022 and a cooperative research fund from the Korean Society of Nephrology (2022).

Conflicts of interest
The authors declare no conflict of interest regarding the present study.

Abbreviations
AKI: Acute kidney injury
ACEi: angiotensin converting enzyme inhibitor
ARB: angiotensin II type 1 receptor blocker
AUC: area under the curve
CMC-CDW: Catholic Medical Center-Clinical Data Warehouse
CKD: chronic kidney disease
CKD-EPI: Chronic Kidney Disease Epidemiology Collaboration
COPD: chronic obstructive pulmonary disease
eGFR: estimated glomerular filtration rate
DNN: deep neural networks
ICU: intensive care units
KDIGO: Kidney Disease: Improving Global Outcomes
LGBM: light gradient boosting machine
MICE: Multiple Imputation by Chained Equations
NSAIDs: non-steroidal anti-inflammatory drugs
OHE: one hot encoded
OR: odds ratio
RASi: renin-angiotensin-aldosterone system inhibitors
ROC: receiver operating characteristic
RRT: renal replacement therapy
sCr: serum creatinine
TRIPOD: transparent reporting of a multivariable prediction model for individual prognosis or diagnosis

Multimedia Appendix 1
ICD-10 codes for comorbid conditions.
Multimedia Appendix 2
Code and Configurations of the Deep Neural Network.
Multimedia Appendix 3
Logistic regression of clinical parameters
Multimedia Appendix 4
Logistic regression of preoperative laboratory parameters

References

1. Thakar CV. Perioperative acute kidney injury. Adv Chronic Kidney Dis. Jan, 2013;20(1):67-75. [doi: 10.1053/j.ackd.2012.10.003] [Medline: 23265598]
2. Biteker M, Dayan A, Tekkeşin A, Can MM, Taycı İ, İlhan E, et al. Incidence, risk factors, and outcomes of perioperative acute kidney injury in noncardiac and nonvascular surgery. Am J Surg. Jan, 2014;207(1):53-59. [doi: 10.1016/j.amjsurg.2013.04.006] [Medline: 24050540]
3. Kim M, Brady JE, Li G. Variations in the risk of acute kidney injury across intraabdominal surgery procedures. Anesth Analg. Nov, 2014;119(5):1121-1132. [doi: 10.1213/ANE.0000000000000425] [Medline: 25191972]
4. Bihorac A, Yavas S, Subbiah S, Hobson CE, Schold JD, Gabrielli A, et al. Long-term risk of mortality and acute kidney injury during hospitalization after major surgery. Ann Surg. May, 2009;249(5):851-858. [doi: 10.1097/SLA.0b013e3181a40a0b] [Medline: 19387314]
5. Park JT. Postoperative acute kidney injury. Korean J Anesthesiol. Jun, 2017;70(3):258-266. [doi: 10.4097/kjae.2017.70.3.258] [Medline: 28580076]
6. Grams ME, Sang Y, Coresh J, Ballew S, Matsushita K, Molnar MZ, et al. Acute kidney injury after major surgery: a retrospective analysis of veterans health administration data. Am J Kidney Dis. Jun, 2016;67(6):872-880. [doi: 10.1053/j.ajkd.2015.07.022] [Medline: 26337133]
7. Boyer N, Eldridge J, Prowle JA-O, Forni LA-O. Postoperative acute kidney injury. Clin J Am Soc Nephrol, 2022;17:1535-1545. [doi: 10.2215/CJN.16541221]
8. Park S, Cho H, Park S, Lee S, Kim K, Yoon HJ, et al. Simple Postoperative AKI Risk (SPARK) classification before noncardiac surgery: a prediction index development study with external validation. J Am Soc Nephrol. Jan, 2019;30(1):170-181. [doi: 10.1681/ASN.2018070757] [Medline: 30563915]
9. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, et al. My surgery risk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg. Apr, 2019;269(4):652-662. [doi: 10.1097/SLA.0000000000002706] [Medline: 29489489]
10. Lei VJ, Luong T, Shan E, Chen X, Neuman MD, Eneanya ND, et al. Risk Stratification for Postoperative Acute Kidney Injury in Major Noncardiac Surgery Using Preoperative and Intraoperative Data. JAMA Netw Open. Dec 2, 2019;2(12):e1916921. [doi: 10.1001/jamanetworkopen.2019.16921] [Medline: 31808922]
11. Vanmassenhove J, Kielstein J, Jörres A, Biesen WV. Management of patients at risk of acute kidney injury. Lancet. May 27, 2017;389(10084):2139-2151. [doi: 10.1016/s0140-6736(17)31329-6] [Medline: 28561005]
12. Kim KJ, Tagkopoulos I. Application of machine learning in rheumatic disease research. Korean J Intern Med. Jul, 2019;34(4):708-722. [doi: 10.3904/kjim.2018.349] [Medline: 30616329]
13. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. Jan 7, 2015;350:g7594. [doi: 10.1136/bmj.g7594] [Medline: 25569120]
14. Kidney Disease Improving Global Outcomes. KDIGO Clinical pactice gideline for acute kidney injury. Kidney Int, 2012;2(1):1-138.
15. Hori D, Katz NM, Fine DM, Ono M, Barodka VM, Lester LC, et al. Defining oliguria during cardiopulmonary bypass and its relationship with cardiac surgery-associated acute kidney injury. Br J Anaesth. Dec, 2016;117(6):733-740. [doi:

10.1093/bja/aew340] [Medline: 27956671]
16. Mizota T, Yamamoto Y, Hamada M, Matsukawa S, Shimizu S, Kai S. Intraoperative oliguria predicts acute kidney injury after major abdominal surgery. Br J Anaesth. Dec 1, 2017;119(6):1127-1134. [doi: 10.1093/bja/aex255] [Medline: 29136086]
17. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, et al. New creatinine- and cystatin c-based equations to estimate GFR without race. N Engl J Med. Nov 4, 2021;385(19):1737-1749. [doi: 10.1056/NEJMoa2102953] [Medline: 34554658]
18. Buuren S, Groothuis-Oudshoorn C. MICE: multivariate imputation by chained equations in R. J Stat Softw. 12/01, 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
19. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. JAMA. 2015 Nov 10, 2015;314(1538-3598 (Electronic)):1966-1967. [doi: 10.1001/jama.2015.15281]
20. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res, 2011;12:2825-2830.
21. Wang L, Han M, Li X, Zhang N, Cheng H. Review of classification methods on unbalanced data sets. IEEE Access, 2021;9:64606-64628. [doi: 10.1109/ACCESS.2021.3074243]
22. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. ArXiv, 2016;abs/1609.06570 [doi: 10.48550/arXiv.1609.06570]
23. Nitesh V. Chawla KWB, Lawrence O. Hall, W. Philip Kegelmeyer. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res, 2002;16:321-357. [doi: 10.48550/arXiv.1106.1813]
24. Lei L, Wang Y, Xue Q, Tong J, Zhou CM, Yang JJ. A comparative study of machine learning algorithms for predicting acute kidney injury after liver cancer resection. PeerJ, 2020;8:e8583. [doi: 10.7717/peerj.8583] [Medline: 32140301]
25. Habehh H, Gohel S. Machine learning in healthcare. Curr Genomics, 2021;22(4):291-300. [doi: 10.2174/1389202922666210705124359] [Medline: 35273459]
26. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al., editors. LightGBM: A highly efficient gradient boosting decision tree. Neural Information Processing Systems; 2017.
27. Chollet F. Keras. GitHub; 2015. URL: https://github.com/fchollet/keras [accessed 2024-05-28]
28. Lewington AJ, Cerdá J, Mehta RL. Raising awareness of acute kidney injury: a global perspective of a silent killer. Kidney Int. Sep, 2013;84(3):457-467. [doi: 10.1038/ki.2013.153] [Medline: 23636171]
29. Susantitaphong P, Cruz DN, Cerda J, Abulfaraj M, Alqahtani F, Koulouridis I, et al. World incidence of AKI: a meta-analysis. Clin J Am Soc Nephrol. Sep, 2013;8(9):1482-1493. [doi: 10.2215/cjn.00710113] [Medline: 23744003]
30. Uchino S, Kellum JA, Bellomo R, Doig GS, Morimatsu H, Morgera S, et al. Acute renal failure in critically ill patients: a multinational, multicenter study. JAMA. Aug 17, 2005;294(7):813-818. [doi: 10.1001/jama.294.7.813] [Medline: 16106006]
31. Kellum JA, Romagnani P, Ashuntantang G, Ronco C, Zarbock A, Anders HJ. Acute kidney injury. Nat Rev Dis Primers. Jul 15, 2021;7(1):52. [doi: 10.1038/s41572-021-00284-z] [Medline: 34267223]
32. Ren Y, Loftus TJ, Datta S, Ruppert MM, Guan Z, Miao 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. May 2, 2022;5(5):e2211973. [doi: 10.1001/jamanetworkopen.2022.11973] [Medline:

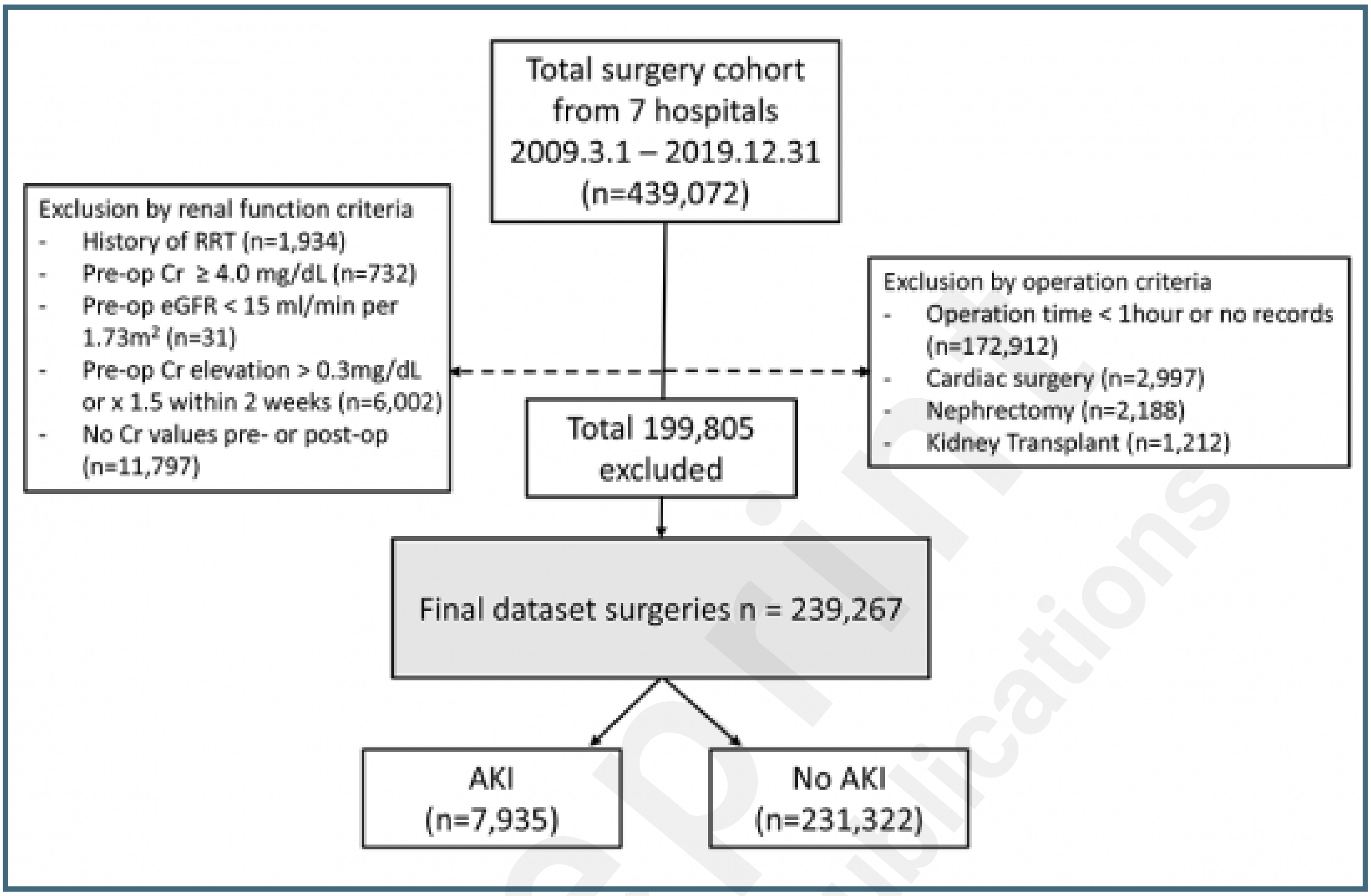
35576007]
33. Sheetz KH, Woodside KA-O, Shahinian VB, Dimick JB, Montgomery JR, Waits SA. Trends in bariatric surgery procedures among patients with ESKD in the United States. Clin J Am Soc Nephrol, 2019;14(8):1193-1199. [doi: 10.2215/CJN.01480219]
34. Dencker EE, Bonde A, Troelsen A, Varadarajan KM, Sillesen M. Postoperative complications: an observational study of trends in the United States from 2012 to 2018. BMC Surg. 2021/11/06, 2021;21(1):393. [doi: 10.1186/s12893-021-01392-z]
35. Kung C-W, Chou Y-H. Acute kidney disease: an overview of the epidemiology, pathophysiology, and management. Kidney Res Clin Pract. 11, 2023;42(6):686-699. [doi: 10.23876/j.krcp.23.001]
36. Polikar R. Ensemble based systems in decision making. IEEE Circuits and Systems Magazine, 2006;6(3):21-45. [doi: 10.1109/MCAS.2006.1688199]
37. Park S, Yi J, Lee YJ, Kwon EJ, Yun G, Jeong JC, et al. Electronic alert outpatient protocol improves the quality of care for the risk of postcontrast acute kidney injury following computed tomography. Kidney Res Clin Pract. Sep, 2023;42(5):606-616. [doi: 10.23876/j.krcp.22.148] [Medline: 37813523]

.

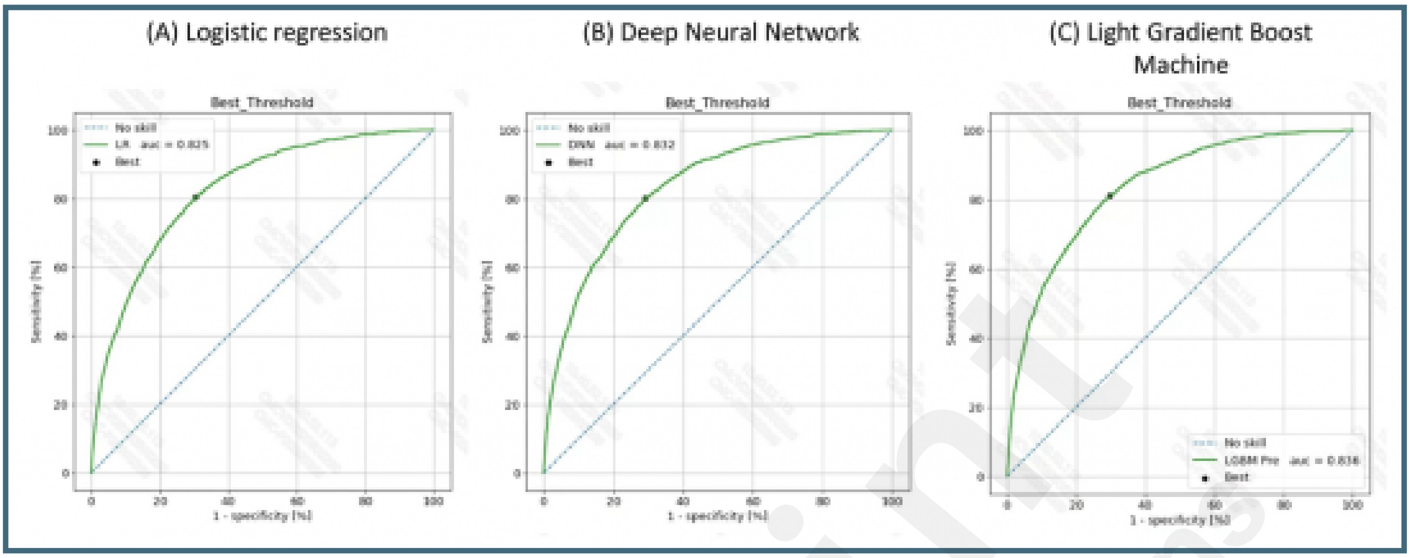
Supplementary Files

Figures

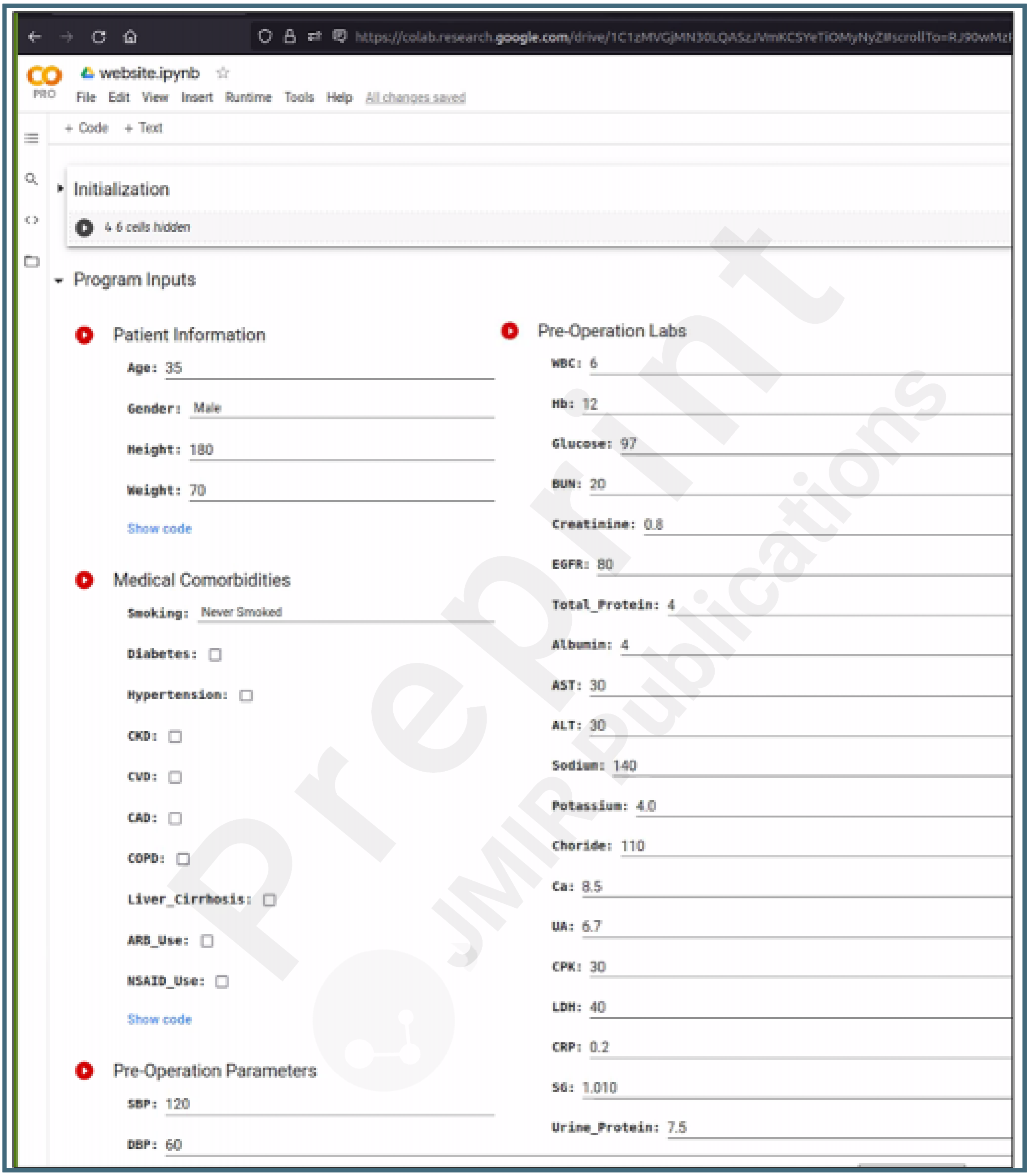
Flow Chart of the Study Population.



AUC of best performance models.



A sample page of the website application.



Multimedia Appendixes

ICD-10 codes for comorbid conditions.
URL: http://asset.jmir.pub/assets/c0fdfa50105195704d3f72b4ac9fdf8b.docx

Code and Configurations of the Deep Neural Network.
URL: http://asset.jmir.pub/assets/3ad5fdeaf6e7945239f766e4bcd6d52d.docx

Logistic regression of clinical parameters.
URL: http://asset.jmir.pub/assets/6d1cff507b55ac7f984ebdfe0dd5a0f1.docx

Logistic regression of preoperative laboratory parameters.
URL: http://asset.jmir.pub/assets/3335e63dbc9413946cfde8dddcd3ee5b.docx