

An explainable machine-learning model for predicting persistent sepsis associated acute kidney injury: development, validation, and comparison with CCL14

Wei Jiang, Yaosheng Zhang, Jiayi Weng, Lin Song, Siqi Liu, Xianghui Li, Siqi Xu, Keran Shi, Luanluan Li, Chuanqing Zhang, Jing Wang, Quan Yuan, Yongwei Zhang, Jun Shao, Jiangquan Yu, Ruiqiang Zheng

Submitted to: Journal of Medical Internet Research
on: June 05, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	32
Figures	33
Figure 1.....	34
Figure 2.....	35
Figure 3.....	36
Figure 4.....	37
Figure 5.....	38
Figure 6.....	39
Figure 7.....	40
Multimedia Appendixes	41
Multimedia Appendix 1.....	42
Multimedia Appendix 2.....	42
Multimedia Appendix 3.....	42
Multimedia Appendix 4.....	42
Multimedia Appendix 5.....	42
Multimedia Appendix 6.....	42
Multimedia Appendix 7.....	42
Multimedia Appendix 8.....	42
Multimedia Appendix 9.....	42

An explainable machine-learning model for predicting persistent sepsis associated acute kidney injury: development, validation, and comparison with CCL14

Wei Jiang^{1*}; Yaosheng Zhang^{2*}; Jiayi Weng^{3*}; Lin Song^{1*}; Siqi Liu^{3*}; Xianghui Li¹; Siqi Xu¹; Keran Shi¹; Luanluan Li¹; Chuanqing Zhang¹; Jing Wang¹; Quan Yuan²; Yongwei Zhang²; Jun Shao¹; Jiangquan Yu^{1*}; Ruiqiang Zheng¹

¹ Department of critical care medicine, Northern Jiangsu People's hospital affiliated Yangzhou University Yangzhou City CN

² School of Clinical and Basic Medicine, Shandong First Medical University & Shandong Academy of Medical Sciences Jinan CN

³ School of Economics and Management, Beijing Jiao Tong University Beijing CN

* these authors contributed equally

Corresponding Author:

Ruiqiang Zheng

Department of critical care medicine, Northern Jiangsu People's hospital affiliated Yangzhou University
No. 98 Nantong West Road
Yangzhou City
CN

Abstract

Background: Persistent sepsis-associated acute kidney injury (SA-AKI) portends worse clinical outcomes and remains a therapeutic challenge for clinicians. Early identification and prediction of persistent SA-AKI is crucial.

Objective: The aim of this study was to develop and validate an interpretable machine learning (ML) model that predicts persistent SA-AKI, and to compare its diagnostic performance with CCL14 in a prospective cohort.

Methods: Four retrospective cohorts and one prospective cohort were used for model derivation and validation. The derivation cohort utilized the MIMIC-IV database, randomly split into 80% for model construction and 20% for internal validation. External validation is conducted using subsets of the MIMIC-III dataset, the e-ICU dataset, and retrospective cohorts from the ICU of a Northern Jiangsu people's hospital. Prospective data from the same ICU were used for validation and compared with urinary CCL14 biomarker measurements. AKI was defined based on serum creatinine and urine output, using the kidney disease: Improving Global Outcomes (KDIGO) criteria. Routine clinical data within the first 24 hours of ICU admission were collected, and eight ML algorithms were utilized to construct the prediction model. Multiple evaluation metrics, including the area under the receiver operating characteristic curve (AUC), were employed to compare predictive performance. Feature importance was ranked using SHAP, and the final model was explained accordingly. In addition, the model is developed into a web-based application using the Streamlit framework to facilitate its clinical application.

Results: In this study, a total of 46,097 sepsis patients from multiple cohorts were enrolled for analysis. Among the 17,928 sepsis patients in the derivation cohort, 8,081 cases (45.1%) developed into persistent SA-AKI. Among eight ML models, the Gradient Boosting Machine (GBM) model demonstrated superior discriminative ability. Following feature importance ranking, a final interpretable GBM model comprising twelve features (AKI stage, ?creatinine, urine output, furosemide dose, BMI, SOFA score, KRT, mechanical ventilation, lactate, Bun, PT and age) was established. The final model accurately predicted the occurrence of persistent SA-AKI in both internal (AUC = 0.870) and external validation cohorts (MIMIC-III subset: AUC = 0.891, e-ICU dataset: AUC = 0.932, North Jiangsu people's Hospital retrospective cohort: AUC = 0.983). In the prospective cohort, the GBM model outperformed urinary CCL14 in predicting persistent SA-AKI (GBM AUC = 0.852 vs. CCL14 AUC = 0.821). Additionally, the model has been transformed into an online clinical tool to facilitate its application in clinical settings.

Conclusions: The interpretable GBM model has been shown to successfully and accurately predict the occurrence of persistent SA-AKI, demonstrating good predictive ability in both internal and external validation cohorts. Furthermore, the model has been demonstrated to outperform the biomarker CCL14 in prospective cohort validation.

(JMIR Preprints 05/06/2024:62932)

DOI: <https://doi.org/10.2196/preprints.62932>

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 visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR publication, my title and abstract will remain visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

An explainable machine-learning model for predicting persistent sepsis associated acute kidney injury: development, validation, and comparison with CCL14

Wei Jiang, PhD^{a#}, Yaosheng Zhang, MD^{b#}, Jiayi Weng, MD^{c#}, Lin Song, PHD^{a#}, Siqi Liu, MD^{c#}, Xianghui Li, MD^a, Siqi Xu, MD^a, Keran Shi, MD^a, Luanluan Li, MD^a, Chuanqing Zhang, MD^a, Jing Wang, PhD^a, Quan Yuan, MD^b, Yongwei Zhang, MD^b, Jun Shao, MD^a, Jiangquan Yu, PhD^{a*} & Ruiqiang Zheng, MD^{a*}

^a Department of Critical Care Medicine, Northern Jiangsu People's Hospital affiliated with Yangzhou University

^b School of Clinical and Basic Medicine, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan 250000, China

^c School of Economics and Management, Beijing Jiao Tong University, Beijing 100044, China

[#] Equally to the contribution

^{*} Corresponding author

Email address zhengruiqiang2021@163.com

Abstract

Background

Persistent sepsis-associated acute kidney injury (SA-AKI) portends worse clinical outcomes and remains a therapeutic challenge for clinicians. Early identification and prediction of persistent SA-AKI is crucial.

Objectives

The aim of this study was to develop and validate an interpretable machine learning (ML) model that predicts persistent SA-AKI, and to compare its diagnostic performance with CCL14 in a prospective cohort.

Methods

Four retrospective cohorts and one prospective cohort were used for model derivation and validation. The derivation cohort utilized the MIMIC-IV database, randomly split into 80% for model construction and 20% for internal validation. External validation is conducted using subsets of the MIMIC-III dataset, the e-ICU dataset, and retrospective cohorts from the ICU of a Northern Jiangsu people's hospital. Prospective data from the same ICU were used for validation and compared with urinary CCL14 biomarker measurements. AKI was defined based on serum creatinine and urine output, using the kidney disease: Improving Global Outcomes (KDIGO) criteria. Routine clinical data within the first 24 hours of ICU admission were collected, and eight ML algorithms were utilized to construct the prediction model. Multiple evaluation metrics, including the area under the receiver operating characteristic curve (AUC), were employed to compare predictive performance. Feature importance was ranked using SHAP, and the final model was explained accordingly. In addition, the model is developed into a web-based application using the Streamlit framework to facilitate its clinical application.

Results

In this study, a total of 46,097 sepsis patients from multiple cohorts were enrolled for analysis. Among the 17,928 sepsis patients in the derivation cohort, 8,081 cases (45.1%) developed into persistent SA-AKI. Among eight ML models, the Gradient Boosting Machine (GBM) model demonstrated superior discriminative ability. Following feature importance ranking, a final interpretable GBM model comprising twelve features (AKI stage, Δ creatinine, urine output, furosemide dose, BMI, SOFA score, KRT, mechanical ventilation, lactate, Bun, PT and age) was established. The final model accurately predicted the occurrence of persistent SA-AKI in both internal (AUC = 0.870) and external validation cohorts (MIMIC-III subset: AUC = 0.891, e-ICU dataset: AUC = 0.932, North Jiangsu people's Hospital retrospective cohort: AUC = 0.983). In the prospective cohort, the GBM model outperformed urinary CCL14 in predicting persistent SA-AKI

(GBM AUC = 0.852 vs. CCL14 AUC = 0.821). Additionally, the model has been transformed into an online clinical tool to facilitate its application in clinical settings.

Conclusions

50 The interpretable GBM model has been shown to successfully and accurately predict the occurrence of persistent SA-AKI, demonstrating good predictive ability in both internal and external validation cohorts. Furthermore, the model has been demonstrated to outperform the biomarker CCL14 in prospective cohort validation.

Keywords: Sepsis; Persistent acute kidney injury; Machine learning; Prediction model; SHAP

55

Introduction

Acute Kidney Injury (AKI) is a common and severe complication in critically ill patients, with sepsis being the most frequent cause[1]. AKI occurring within 7 days after the onset of sepsis is defined as sepsis-associated acute kidney injury (SA-AKI) [2]. Studies have estimated that 68% of
60 patients with sepsis present with AKI on admission, 40% present with severe AKI, and 27% require subsequent kidney replacement therapy (KRT) during their intensive care unit (ICU) stay [3]. The development of SA-AKI correlates significantly with higher mortality rates and an increased risk of chronic kidney disease [4, 5].

SA-AKI is a complex clinical syndrome, depending on interactions between genotypes and
65 exposures, SA-AKI can lead to various clinical phenotypes. This heterogeneity complicates efficacy assessments in clinical trials for sepsis interventions because different treatments may only be beneficial for specific disease subtypes [2]. Recently, the Acute Disease Quality Initiative (ADQI) 16 Workgroup suggested that AKI be classified as transient (a complete reversal of AKI within 48 h) or persistent (the continuance of AKI beyond 48 h) [6]. Compared to transient AKIs, persistent ones are
70 associated with ongoing host response dysregulation and adverse outcomes in critically ill septic patients [7].

Early identification of persistent AKI has significant implications for risk stratification and individualized treatment. Firstly, the duration of an AKI closely relates to patient prognosis and risks of end-stage renal failure. Recent evidence shows that two-thirds of all cases recover kidney function
75 within three to seven days, while those with persistent conditions show significantly reduced survival rates after one year [8]. Additionally, continued presence also increases individuals' risks for chronic kidney disease (CKD)[9]. Early recognition and proactive intervention could potentially affect progression towards CKD. Secondly, the duration of AKI is closely related to the need for kidney replacement therapy (KRT). Studies show that some patients may benefit from starting KRT earlier,
80 while others might not require it at all, as they can quickly recover kidney function [10]. Therefore, predicting the short-term reversibility of AKI could help assess potential needs for KRT and determine optimal timing for its initiation.

However, due to complex mechanisms potentially co-existing in SA-AKI pathophysiology including circulatory dysfunction, inflammatory response, mitochondrial dysfunction and metabolic reprogramming[1], single-method diagnoses of persistent SA-AKI have limited effectiveness.
85 Traditional urinary indicators and renal ultrasound markers reportedly play a limited role in distinguishing between transient and persistent AKIs[11-14]. Furthermore, most biomarkers assessing functional or injury value perform poorly in predicting persistent AKI[15-18] while others still require further clinical validation[19, 20] with none currently approved for clinical use. There
90 remains a lack of effective methods on how to rapidly identify potential cases of persistent SA-AKI.

In recent years, machine learning (ML) methods derived from Electronic Medical Records (EMR) have gained attention and recognition among clinicians. The widespread application of EMRs within hospitals has made patient data collection more accurate and convenient. Currently, some

studies suggest that predictive models built using ML methods are promising assessment tools for AKI[21-26]. Although these models demonstrate high diagnostic performance, they do not differentiate between clinical subtypes, which limits their guidance on personalized treatment strategies. Additionally, ML models remain difficult to interpret directly, presenting what's known as a "black box" problem, limiting their application by clinicians. Jiang's study used an ML method to build a predictive model about this subtype but did not target septic patients, nor was the model externally validated [27].

In our study, we developed an interpretable ML model in four retrospective cohorts and one prospective cohort, aimed at early and accurate prediction of persistent sepsis-associated acute kidney injury. We elucidated the importance of characterization and interpreted the model using the SHAP method while comparing it to an existing biomarker, CCL14, which has high diagnostic performance.

Methods

Data source and study population

The derivation cohort and internal validation cohort originate from the MIMIC-IV database. The MIMIC-IV database is a publicly available multi-parameter intensive care database provided by the Massachusetts Institute of Technology (MIT)[28]. It includes critically ill patients admitted to the ICU at Beth Israel Deaconess Medical Center in Boston, Massachusetts, from 2008 to 2019.

The datasets used for external validation include the MIMIC-III dataset, e-ICU dataset, and two cohorts admitted to the ICU of Northern Jiangsu People's Hospital. The external validation cohort 1 is a subset of the MIMIC-III dataset; the MIMIC-III database is a publicly available multi-parameter intensive care database provided by the Massachusetts Institute of Technology (MIT). It includes critically ill patients admitted to the ICU at Beth Israel Deaconess Medical Center in Boston, Massachusetts, from 2008 to 2012[29]. Considering that the MIMIC-III dataset and the MIMIC-IV dataset have overlapping sections, we only included the population from this dataset between 2001 and 2008. The external validation cohort 2 consists of sepsis patients collected from the e-ICU database. The e-ICU Collaborative Research Database is a multi-center database comprising deidentified health data associated with over 200,000 admissions to ICUs across the United States between 2014 and 2015. The database includes vital sign measurements, care plan documentation, severity of illness measures, diagnosis information, and treatment information. Data is collected through the Philips e-ICU program, a critical care telehealth program that delivers information to caregivers at the bedside [30]. The external validation cohort 3 is a retrospective cohort of patients admitted to the ICU at Northern Jiangsu People's Hospital from October 2021 to January 2023. The external validation cohort 4 is a prospective cohort of patients admitted to the ICU at Northern Jiangsu People's Hospital from October 2023 to March 2024, where urine samples were collected upon their admission and CCL14 levels were measured using ELISA kits.

Both development and validation cohorts adopted identical inclusion and exclusion criteria; we included adult patients who developed sepsis after their first admission to the ICU, while excluding those with stage 4–5 chronic kidney disease.

Ethical considerations

Some of the data in this study is based on a third-party anonymous public database, for which we have already obtained approval from the Institutional Review Board. Therefore, this portion of data does not require ethical review. In order to access this database, we have completed an online training course and examination for protecting human research participants provided by the U.S. National Institutes of Health (ID number 54780440). The data collected by the ICU at Northern Jiangsu People's Hospital has been approved by its Ethics Committee (ID number: 2021ky230) and registered on the China Clinical Trial Registration Platform (Registration number:

ChiCTR2100050540).

Primary outcomes and definitions

The study outcome was the occurrence of persistent SA-AKI. Persistent SA-AKI is defined as developing into SA-AKI upon admission to the ICU, and if the duration of SA-AKI exceeds 48 hours. Patients who die or start KRT within 48 hours after developing SA-AKI are also considered to have persistent sepsis-associated AKI[6].

According to the definition of SEPSIS-3, we identified patients with confirmed or suspected infection and a Sequential Organ Failure Assessment (SOFA) score increase of two or more [31]. We determined the clinician's recognition of suspected infection through two simultaneous events in the electronic health records: 1) prescription of antibiotics, and 2) ordering specific fluid cultures. These two events need to occur within a specific time frame and have the following options: In option 1, fluid culture is performed first, and antibiotic use must be initiated within 72 hours. In option 2, antibiotic dosing is administered first, and fluid culture must be completed within 24 hours. We excluded all antibiotics given as a single dose in the operating room. We also excluded antibiotics that were not accompanied by fluid cultures. We assumed a SOFA score of zero prior to ICU admission. If individual components of SOFA were missing, no contribution was made to the total score [32]. The daily total SOFA scores were calculated, and an increase of two points within 24 hours was considered abnormal [33]. Considering difficulties in interpreting neurological SOFA when sedation therapy is being concurrently administered, it was not included in the overall scoring category [34].

The diagnostic criteria for AKI follow the standards of Kidney Disease: Improving Global Outcomes (KDIGO): an increase in serum creatinine (Scr) exceeding $26.5\mu\text{mol/L}$ (0.3mg/dl) within 48 hours; an increase in Scr by more than 50% from baseline, lasting for 7 days; urine output less than $0.5\text{ml}/(\text{kg}\cdot\text{h})$, lasting for more than 6 hours [35]. Baseline Scr was defined as the lowest Scr value during 7 days before ICU admission [36, 37]. When pre-admission Scr was not available, baseline creatinine was estimated using the MDRD equation by back calculation from an estimated glomerular filtration rate of $75\text{ mL}/\text{min}/1.73\text{ m}^2$ [38]. Either urine-based or creatinine-based criteria, or a combination of both, are used to determine if a patient meets KDIGO AKI criteria.

After determining sepsis and acute kidney injury separately, we applied the definition of SA-AKI from the ADQI 28 working group. We compared the diagnosis day of sepsis with the diagnosis day of AKI. If AKI occurs within 1 to 7 days after the diagnosis of sepsis, patients are classified as SA-AKI according to ADQI criteria [2]. If AKI occurs before sepsis, patients do not meet the definition of SA-AKI.

Data collection and processing

We utilized demographic characteristics, vital sign measurements, and laboratory data collected within the first day of ICU admission to identify features and construct a predictive model (Multimedia Appendix 1). All data were sourced from electronic medical records. Due to potential multicollinearity among features that could affect prediction accuracy, any feature with a high correlation (correlation coefficient > 0.6) with another in Spearman's correlation analysis was removed if it had less relevance to the outcome (Multimedia Appendix 2). Additionally, we excluded any feature missing more than 30% of its values in subsequent analyses to minimize bias due to missing data. To normalize the data, we use the StandardScaler z-score normalization function. To process categorical features, we used label encoding. Given the large number of categorical variables in our dataset, label encoding was chosen over one-hot encoding to mitigate the dimensionality expansion problem and thereby improve computational efficiency.

Finally, we developed our predictive model using 58 features including age, gender, BMI, maximum heart rate(HR), minimum mean arterial pressure(MAP), maximum respiratory rate(RR),

highest temperature, maximum blood glucose level, AKI stage, 24-hour creatinine change, past medical history (hypertension, diabetes, coronary heart disease, chronic kidney disease, chronic heart failure, chronic liver disease chronic cardiac diseases and chronic obstructive pulmonary disease), source of infection (pulmonary infection, blood infection, urinary tract infections, abdominal infections, skin infections and others), interventional treatment (use of mechanical ventilation(MV), kidney replacement therapy(KRT), diuretics, statins, ACEI/ARB, aminoglycoside antibiotics, glycopeptide antibiotics, Nonsteroidal Anti-inflammatory Drugs, acyclovir antiviral drugs), furosemide dosage, urine output, maximum pH value, maximum lactate value, 24-hour lactate clearance rate, minimum arterial oxygen tension value(PaO₂), 24-hour oxygen tension change value, maximum carbon dioxide partial pressure(PCO₂), 24-hour carbon dioxide partial pressure change, maximum base excess(BE), maximum white blood cell count(WBC), maximum platelet count, minimum red blood cell count(RBC), mean corpuscular volume(MCV), mean corpuscular hemoglobin concentration(MCHC), bicarbonate (HCO₃⁻), maximum potassium(K⁺), maximum sodium (Na⁺), maximum chloride (Cl⁻), maximum calcium (Ca⁺), maximum activated partial thromboplastin time(APTT), maximum prothrombin time(PT), maximum blood urea nitrogen (BUN), Glasgow score(GCS), SOFA score.

For the prospective cohort, patients suspected of sepsis or diagnosed with sepsis upon ICU admission had 10 ml of urine collected. The sample was then centrifuged at 3000 rpm for 5 minutes and left to stand for 30 minutes before the supernatant was collected and stored at ≤-70°C for subsequent analysis. Urinary CCL14 levels were measured using an enzyme-linked immunosorbent assay kit (Conlon Bioproducts; detection instrument: Rayto RT-6100 microplate reader). Operators were blinded to the experimental research content and strictly adhered to the instructions provided with the kit.

Model development and comparison

The development cohort originates from MIMIC-IV, which is randomly divided into an 80% training cohort and 20% validation cohort (internal validation). Ten-fold cross-validation is used for the 80% training set to ensure generalizability. In addition, an external dataset is used for testing (external validation).

The predictive model was developed using 58 features included in the data. Missing data were handled using multiple imputation [39], and the proportion of missing data for each variable can be found in Multimedia Appendix 1. Eight machine learning models were used to predict persistent SA-AKI, including Artificial Neural Networks (ANN), Decision Trees (DT), Gradient Boosting Machines (GBM), K-Nearest Neighbors Enhancer (KNN), Logistic Regression (LR), Catboost, Support Vector Machines (SVM) and Extreme Gradient Boosting (XGboost). To optimize prediction models, Bayesian optimization combined with manual fine-tuning was employed to obtain final hyperparameters.

To assess the reliability of these models, several commonly used evaluation metrics were utilized, such as Area Under Receiver Operating Characteristic Curve (AUC), sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), accuracy and F1 score.

Feature selection and model explanation

Interpreting ML models correctly often poses challenges. The SHAP method is a technique that can rank the importance of input features and explain the results of predictive models, commonly used to overcome the "black box" problem in ML[39]. By using SHAP values to assist feature selection, we reduced the predictive model from 58 features down to three based on their importance ranking. We then selected for further analysis the final model with optimal predictive ability during this reduction process. Using DeLong's non-parametric method, we compared differences between AUCs and gradually decreased the number of features in our chosen ML model until there was a

235 significant drop in AUC.

The SHAP method provides both global and local interpretations for models. Global interpretation offers consistent and accurate attribution values for each feature within a model, demonstrating correlations between input features and persistent SA-AKI. Local interpretation allows specific predictions for individual populations by entering particular data.

240 Webpage deployment tool based on streamlit framework

To facilitate the application of this model in a clinical setting, the final prediction model has been implemented into a web application built on a Python framework based on streamlit. When provided with corresponding feature values in the final model, the application can return the probability of persistent SA-AKI.

245

Statistical analysis

Data analysis was conducted using Python 3.9.18 and R 4.2.2. Continuous variables with a non-normal distribution were represented by the median of the interquartile range and compared using the Mann-Whitney U test or Kruskal-Wallis H test. Categorical variables were expressed in percentages and compared using the chi-square test or Fisher's exact test. Analysis of covariance (ANCOVA) was used to adjust for confounding factors. We utilized AUC to assess predictive ability and established optimal cut-off values by maximizing Youden's index (sensitivity + specificity -1). A two-tailed P value <0.05 was considered statistically significant.

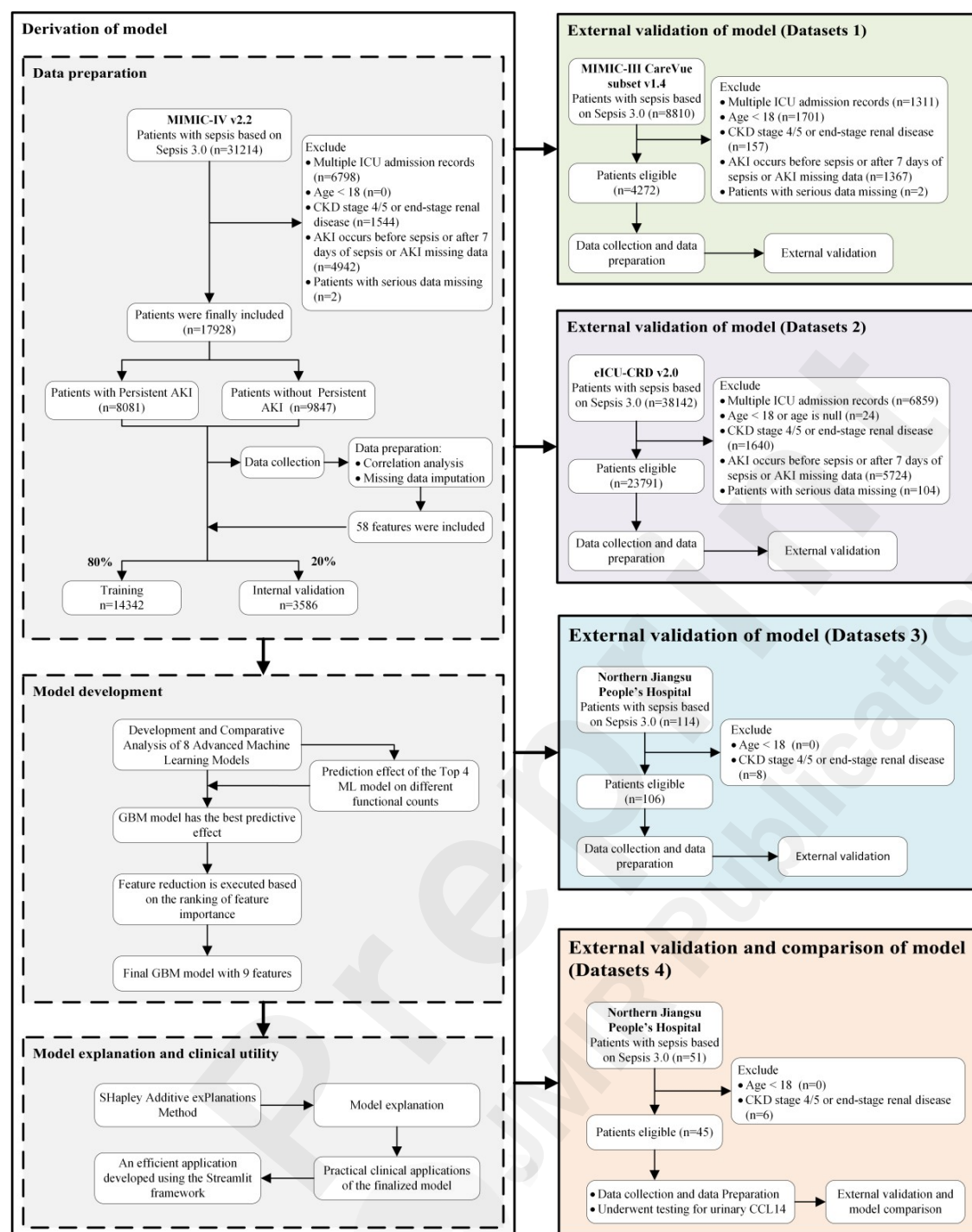
250

Results

255 Patient characteristics

A total of 17,928 sepsis patients were included in the MIMIC-IV cohort for predictive model construction. According to Sepsis 3.0 criteria, a total of 31,214 septic patients were included in the MIMIC-IV dataset. After excluding 6,798 cases with repeated ICU admissions, 1,544 CKD4-5 stage cases, and 4,972 cases where AKI occurred before or after seven days of sepsis onset, as well as two severe data-missing cases, a total of 17,928 patients were analyzed including persistent SA-AKI (8,081) and non-persistent SA-AKI (9,847). Out of these patients, 14,342 (80%) were allocated to the validation cohort, while the remaining 3,586 (20%) went into the internal validation set. Moreover, in external validation cohorts, MIMIC-III subset, e-ICU datasets similar to MIMIC-IV dataset after excluding multiple ICU admissions, CKD4-5 stages, severe data missing and AKI occurring before or after seven days from sepsis onset, 4272 and 23791 patients, respectively, got enrolled. The Northern Jiangsu People's Hospital retrospective cohort included 114 septic patients, out of which eight CKD4-5 stage patients got excluded, leaving behind 106 patients. Furthermore, a prospective cohort at Northern Jiangsu People's Hospital collected 51 cases, out of which 6 patients with CKD4-5 got excluded, resulting in 45 cases for analysis. Comparison between training set, internal verification set, and external verification sets regarding demographic statistics and clinical variables can be seen in Multimedia Appendix 3. Detailed information about research design is available in Fig.1.

270



275 Figure 1 Study flow chart. Sepsis-associated acute kidney injury; CKD: chronic kidney disease; GBM: Gradient Boosting Machine; ICU: intensive care unit; SHAP: SHapley Additive explanation; CCL14:C-C motif chemokine ligand 14.

280 Out of the 17,928 septic patients in the derived cohort, 8,081 (45.1%) developed into persistent SA-AKI, including 2,299 who met AKI Stage 1 criteria, 4,545 who met AKI Stage 2 criteria, and 1,237 who met AKI Stage 3 criteria. The demographic and clinical characteristics between non-persistent SA-AKI and persistent SA-AKI can be found in Table 1.

Table 1: Comparison of demographic and clinical characteristics and outcomes between Persistent SA-AKI and Non-Persistent SA-AKI in the derivation cohort.

	Non-Persistent SA-AKI	Persistent SA-AKI	p. overall
Variables	N=9847	N=8081	

AKI stage			.000
0	5421 (55.1%)	0 (0.00%)	
1	2161 (21.9%)	2299 (28.4%)	
2	2117 (21.5%)	4545 (56.2%)	
3	148 (1.50%)	1237 (15.3%)	
BMI, kg/m ²	26.6 [23.2;30.5]	28.9 [25.0;34.0]	<.001
Gender, Male, n%	5658 (57.5%)	4741 (58.7%)	0.106
<i>Comorbidities, n%</i>			
Hypertension	4363 (44.3%)	3873 (47.9%)	<.001
Diabetes	2547 (25.9%)	2486 (30.8%)	<.001
Coronary Atherosclerosis	2598 (26.4%)	2864 (35.4%)	<.001
Chronic Kidney Disease	1307 (13.3%)	1398 (17.3%)	<.001
Chronic Heart Failure	2014 (20.5%)	2650 (32.8%)	<.001
Chronic Liver Disease	815 (8.28%)	851 (10.5%)	<.001
COPD	1354 (13.8%)	1351 (16.7%)	<.001
<i>Infection sources, n%</i>			
Lung infection	3575 (36.3%)	3487 (43.2%)	<.001
Intestinal infection	1284 (13.0%)	1078 (13.3%)	.569
Catheter related infection	475 (4.82%)	688 (8.51%)	<.001
Urinary system infection	1814 (18.4%)	1507 (18.6%)	.712
Skin infection	468 (4.75%)	375 (4.64%)	.751
<i>Severity scale</i>			
GCS	15.0 [13.0;15.0]	15.0 [13.0;15.0]	<.001
SOFA	4.00 [3.00;6.00]	6.00 [4.00;9.00]	<.001
<i>Vital signs</i>			
HR_max, beats/min	103 [91.0;117]	104 [91.0;119]	.010
MAP_min, mmHg	59.0 [52.0;65.0]	57.0 [50.0;62.0]	<.001
RR_max, times/min	27.0 [24.0;31.0]	28.0 [24.0;32.0]	<.001
Temp_max, °C	37.4 [37.0;37.9]	37.4 [37.0;37.9]	.014
<i>Laboratory findings</i>			
PaO ₂ _min, mmHg	84.8 [56.0;109]	75.0 [46.0;100]	<.001
ΔPaO ₂ , mmHg	-56.00 [-128.42; -10.98]	-42.00 [-146.00;5.00]	<.001
PCO ₂ _max, mmHg	44.8 [40.0;50.0]	47.0 [42.0;54.0]	<.001

ΔPCO_2 , mmHg	-0.57 [-2.43;1.21]	-1.00 [-4.77;2.00]	<.001
PH_min	7.34 [7.29;7.38]	7.32 [7.25;7.37]	<.001
Lactate_max, mmol/L	2.01 [1.40;2.90]	2.40 [1.60;3.60]	<.001
Lactate clean rate	0.10 [0.01;0.19]	0.08 [-0.03;0.22]	<.001
BE_max, mmol/L	0.00 [-1.00;2.00]	0.13 [-1.00;3.00]	.003
Glucose_max, mg/dL	160 [127;202]	173 [140;220]	<.001
BUN_max, mmol/L	19.0 [13.0;29.0]	23.0 [16.0;34.0]	<.001
WBC_max, k/uL	13.3 [9.50;17.8]	14.3 [10.3;19.4]	<.001
Platelets_max, k/uL	200 [146;270]	197 [144;267]	.041
RBC_min, k/uL	3.33 [2.85;3.82]	3.23 [2.76;3.78]	<.001
MCHC_min, g/dL	32.9 [31.8;33.9]	32.5 [31.4;33.6]	<.001
MCV_min, fL	90.0 [86.0;94.0]	90.0 [86.0;94.0]	<.001
RDW_max, %	14.5 [13.5;15.9]	14.8 [13.8;16.5]	<.001
Calcium_max, mEq/L	8.40 [8.00;8.90]	8.45 [8.10;8.90]	.002
Potassium_max, mEq/L	4.40 [4.00;4.80]	4.50 [4.10;5.00]	<.001
Chloride_max, mEq/L	107 [104;111]	107 [103;111]	.217
Aniongap_max, mmol/L	15.0 [13.0;18.0]	16.0 [13.0;19.0]	<.001
PT_max, seconds	14.5 [12.9;16.7]	15.5 [13.5;18.9]	<.001
APTT_max, seconds	31.9 [28.0;38.4]	34.6 [29.4;46.6]	<.001
<i>Interventions, n%</i>			
KRT	48 (0.49%)	522 (6.46%)	<.001
MV	4492 (45.6%)	5263 (65.1%)	<.001
Norepinephrine	1806 (18.3%)	2406 (29.8%)	<.001
Diuretic	2556 (26.0%)	3053 (37.8%)	<.001
Furosemide dose, mg	0.00 [0.00;0.00]	0.00 [0.00;20.0]	<.001
Statin	3258 (33.1%)	3166 (39.2%)	<.001
ACEI/ARBS	866 (8.79%)	758 (9.38%)	.183
Aminoglycoside	275 (2.79%)	252 (3.12%)	.215
Glycopeptide:	5355 (54.4%)	4672 (57.8%)	<.001
NSAIDS	3643 (37.0%)	3710 (45.9%)	<.001
Acyclovir	457 (4.64%)	264 (3.27%)	<.001
Fluid balance	2333 [470;4744]	3547 [1131;6362]	<.001
Urine output	1910 [1300;2754]	1345 [805;2015]	<.001
<i>Outcomes</i>			

Length of stay, days	1.99 [1.23;3.26]	3.79 [1.86;7.18]	.000
ICU mortality, n%	277 (2.81%)	1326 (16.4%)	<.001

SA-AKI: sepsis associated acute kidney injury; BMI: body mass index; HR: MAP: mean arterial pressure; Temp: temperature; COPD: chronic obstructive pulmonary disease; SOFA: Sequential Organ Failure Assessment; GCS: Glasgow Coma Scale; KRT: kidney replacement therapy; MV: mechanical ventilation; ACEI/ARBs: angiotensin converting enzyme inhibitor/angiotensin receptor blocker; NSAIDs: Nonsteroidal Anti-inflammatory Drugs; PH: potential of hydrogen; PaCO₂: partial pressure of carbon dioxide; PaO₂: partial pressure of oxygen; BE: base excess; WBC: white blood cell; APTT: activated partial thromboplastin time; BUN: blood urea nitrogen; MCHC: mean corpuscular hemoglobin concentration; MCV: mean corpuscular volume; RDW: Red blood cell distribution width; PT: Prothrombin time; ICU: intensive care unit.

Model development and performance comparison

Eight ML models were generated using clinical data collected within 24 hours after admission to the ICU in order to predict whether septic patients would develop persistent SA-AKI during this period. Among these eight models, the GBM model (AUC = 0.872) had the best predictive performance for Persistent SA-AKI, followed by the CatBoost model (AUC = 0.870) and XGBoost model (AUC = 0.859). The discriminative abilities of these eight models can be seen in Multimedia Appendix 4, while ROC curves and SHAP summary plots for the top 12 features of the four best-performing ML models are shown in Fig. 2A and Multimedia Appendix 5 respectively.

During feature reduction based on feature importance rankings, changes in AUC among four types of models indicated that GBM maintained almost optimal predictive ability among all four types of models, as shown in Fig. 2B. Therefore, it was observed that among the aforementioned four models, GBM performed best at predicting Persistent SA-AKI. The performance of GBM with different numbers of features is illustrated in Figure 2C and Multimedia Appendix 6.

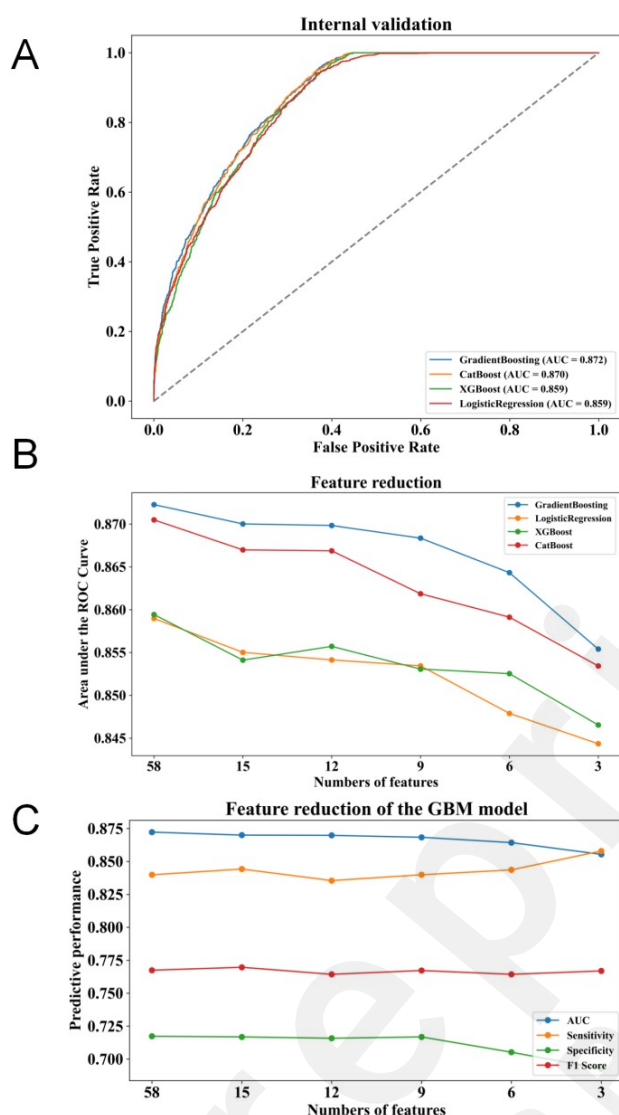


Figure 2: Performance of ML models to predict Persistent SA-AKI. (A) ROC curves of the top four best-performing ML models. (B) AUCs of the top four best performing ML models with varied numbers of features. (C) AUC, sensitivity, specificity, and F1 score of the GBM model with varied numbers of features. SA-AKI: sepsis-associated acute kidney injury; AUC: area under the ROC curve; ROC: receiver-operating characteristic; GBM: gradient boosting machine; XGboost: eXtreme gradient boosting; CatBoost: categorical boosting; ML: machine learning.

Identification of the final model

The final model was determined during the feature reduction process of the GBM model. As shown in Fig. 2C and Multimedia Appendix 7A, the 58-feature model significantly outperformed the 3-feature model ($\Delta\text{AUC} = 0.017$, $P = 0.000$), the 6-feature model ($\Delta\text{AUC} = 0.008$, $P = 0.000$) and the 9-feature model ($\Delta\text{AUC} = 0.004$, $P = 0.028$) in predicting whether ICU patients would develop persistent SA-AKI, but it did not have a significant advantage over the 12-feature model ($\Delta\text{AUC} = 0.002$, $P = 0.130$). Compared with the 58-feature model, the 12-feature one showed better net benefits at higher threshold probabilities. Meanwhile, its area under the P-R curve was only slightly lower than that of the 58 feature one, indicating both models had similar and high clinical utility as shown in Multimedia Appendix 7B-F. Therefore, we chose to use a GBM Model based on AKI stage, Δ Creatinine, Urine output, Furosemide dose, BMI, SOFA score, KRT, Mechanical ventilation, BUN, lactate, PT and Age - these 12 features for further analysis. Finally, the AUC of our ultimate GBM Model for predicting Persistent SA-AKI is found to be at an impressive level of 0.870 (95%CI (0.859-0.870)), with sensitivity being 0.836, specificity being 0.716, PPV being 0.704, NPV being 0.843, accuracy rate standing at 0.769, and F1 Score is calculated to be around 0.764. In addition, further cross-validation was performed to verify the appropriate sample size for this study

and the robustness of the model to locus variation. As shown in Multimedia Appendix 8, the final model obtained mean aucs of 0.883 ± 0.009 and 0.884 ± 0.008 for 5-fold and 10-fold cross-validation, respectively.

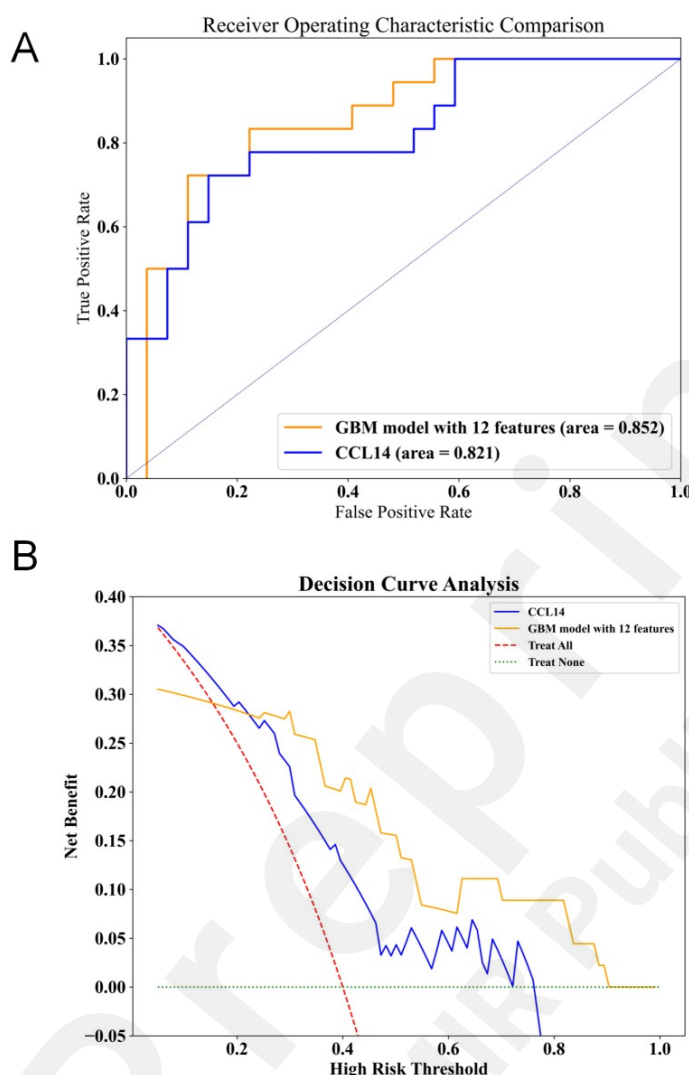


Figure 3 Comparison between the final GBM model and CCL14. (A-B) ROC curves (A) and DCA curves (B) of CCL14 and the final GBM model with 8 features. These plots represented the predictive performance in the prospective external validation cohort. AUC: area under the ROC curve; DCA: decision curve analysis; ROC: receiver-operating characteristic; GBM: gradient boosting machine; CCL14: C-C motif chemokine ligand 14;

External validation of the final mode

For external validation, the final model achieved an AUC of 0.891 on a subset of the MIMIC-III dataset, an AUC of 0.932 on the e-ICU dataset, and an AUC of 0.983 in a retrospective cohort from Northern Jiangsu People's Hospital. These results are similar to those obtained during internal validation (AUC = 0.870) (Table 2), indicating that our final model demonstrated excellent diagnostic performance in both internal and external validations. We further compared our final model with biomarkers known for their high accuracy in predicting persistent SA-AKI within a prospective cohort study. The results suggest that our model's AUC is higher than CCL14 ($\Delta\text{AUC} = 0.031$) (Fig. 3A). The DCA curve also shows that our final model has greater clinical utility compared to CCL14 (Fig. 3B).

Table 2 Performance of the GBM model with different External validations for Persistent SA-AKI prediction

Dataset	AUR OC	AUPR C	Accurac y	F1 Score	Precision	Recall	Specificity	NPV	Youden's
e-ICU	0.932	0.795	0.850	0.722	0.692	0.754	0.884	0.912	0.637
Retrospective cohort of Northern Jiangsu People's Hospital	0.983	0.981	0.909	0.909	0.833	1.000	0.833	1.000	0.833
Prospective cohort of Northern Jiangsu People's Hospital	0.852	0.742	0.822	0.765	0.813	0.722	0.889	0.828	0.611
MIMIC-III	0.891	0.848	0.802	0.794	0.744	0.851	0.763	0.863	0.614

SA-AKI: sepsis-associated acute kidney injury; AUC: area under the receiver-operating characteristic curve; NPV: negative predictive value; GBM: gradient boosting machine, ICU: intensive care unit.

350 Model explanation

Given that clinicians often struggle to accept a predictive model they cannot directly interpret and understand, we employed the SHAP method to explain the output of our final model. This was achieved by calculating each variable's contribution to the prediction. This interpretable approach provides two types of explanations: global interpretations at the feature level and local interpretations at an individual level. Global interpretation describes the overall function of the model. As shown in SHAP summary plots (Fig. 4A and B), features' contributions to models are evaluated using average SHAP values and displayed in descending order. Additionally, SHAP dependence plots help understand how individual features impact predictions from our model. Fig 5 shows a comparison between actual values for these 12 features and their corresponding SHAP values, where a positive value indicates a higher risk for persistent SA-AKI within our model - or put differently, it pushes decisions towards "persistent SA-AKI". For instance, AKI stages greater than or equal to one or creatinine changes over 24 hours greater than or equal to 0.1 mg/dl, as well as use of furosemide on admission day, renal replacement therapy (RRT) and septic patients with KRT all have positive SHAP values pushing decisions towards "persistent SA-AKI". Similarly, SOFA scores above eight points, urine outputs less than 1380 ml/24 h, BMI indices above 24.92, age older than 66 years also push decision-making towards this category.

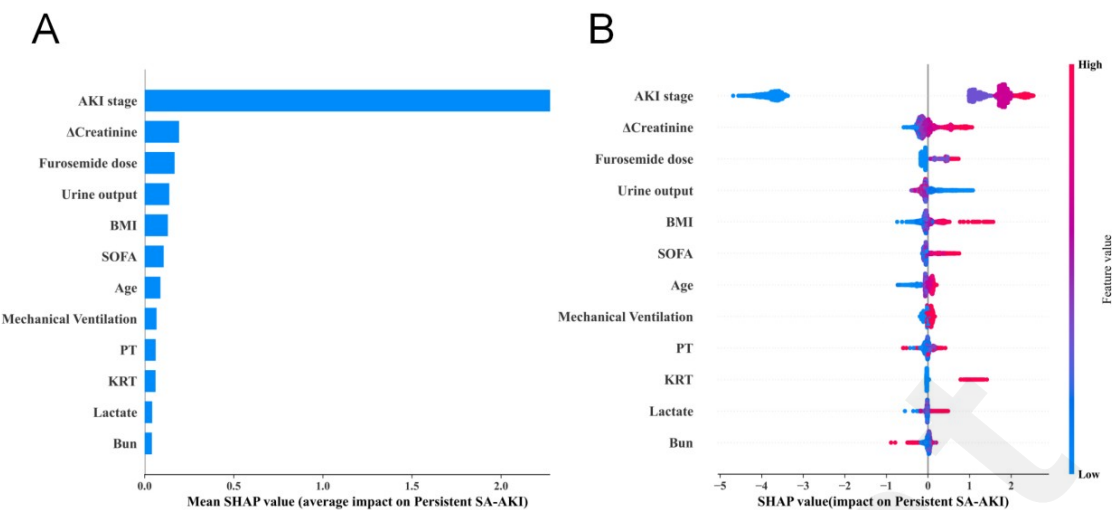


Figure 4: Global model explanation by the SHAP method. (A) SHAP summary bar plot. (B) SHAP summary dot plot. The probability of Persistent SA-AKI development increases with the SHAP value of a feature. A dot is made for SHAP value in the model for each single patient, so each patient has one dot on the line for each feature. The colors of the dots demonstrate the actual values of the features for each patient, as red means a higher feature value and blue means a lower feature value. The dots are stacked vertically to show density. SA-AKI:sepsis-associated acute kidney injury;AKI:acute kidney injury; ΔCreatinine:Changes in creatinine within 24 hours; BMI: body mass Index; SOFA: sequential organ failure assessment; KRT: kidney renal therapy;PT:Prothrombin time.

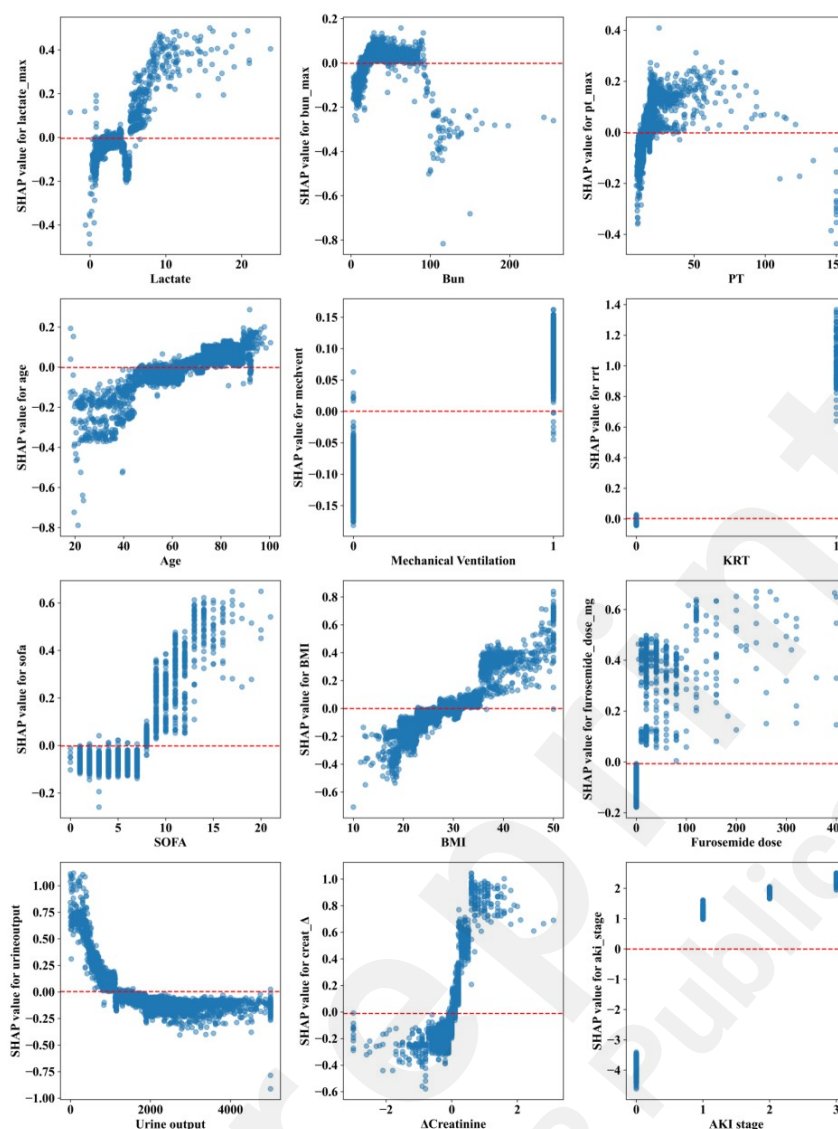


Figure 5: SHAP dependence plot. Each dependence plot shows how a single feature affects the output of the prediction model, and each dot represents a single patient. For example, SOFA scores above eight points or urine outputs less than 1380 ml/24 h or BMI indices above 24.92 or age older than 66 years push the decision towards the “Persistent SA-AKI” class. SHAP values are represented by the y-axis, and actual values are represented by the x-axis. The SHAP values for specific features exceeding zero push the decision towards the “Persistent SA-AKI” class. SA-AKI: sepsis-associated acute kidney injury; AKI: acute kidney injury; Δ Creatinine: Changes in creatinine within 24 hours; BMI: body mass Index; SOFA: sequential organ failure assessment; KRT: kidney renal therapy; PT: Prothrombin time.

Furthermore, local interpretation analyses how specific predictions are made for particular individuals, including personalized input data. Fig. 6A presents a patient who developed persistent SA-AKI during their ICU stay. According to our predictive model, this patient is categorized under “persistent SA-AKI” with probability being at 87%. The figure reveals that factors such as AKI stage, BMI, Urine output, Δ Creatinine and age pushed decision-making toward persistent SA-AKI. while lower SOFA scores, absence of mechanical ventilation, RRT, and diuretics reduced the risk of AKI. Conversely, these factors could increase the patient's risk for persistent SA-AKI.

In Figure 6B, we observe a patient who did not develop persistent SA-AKI during their ICU stay. Figure 6B displays features and their actual measurements that push decision-making towards “non-persistent SA-AKI”. The decision in this case leans towards “non-persistent SA-AKI” with a probability of 0.47%. Additionally, Figure 6C shows an explanatory power chart for patients in our internal validation cohort. The x-axis represents each patient, while the y-axis indicates feature contributions. Each individual patient has increased red portions, indicating higher probabilities for “persistent SA-AKI” decisions.

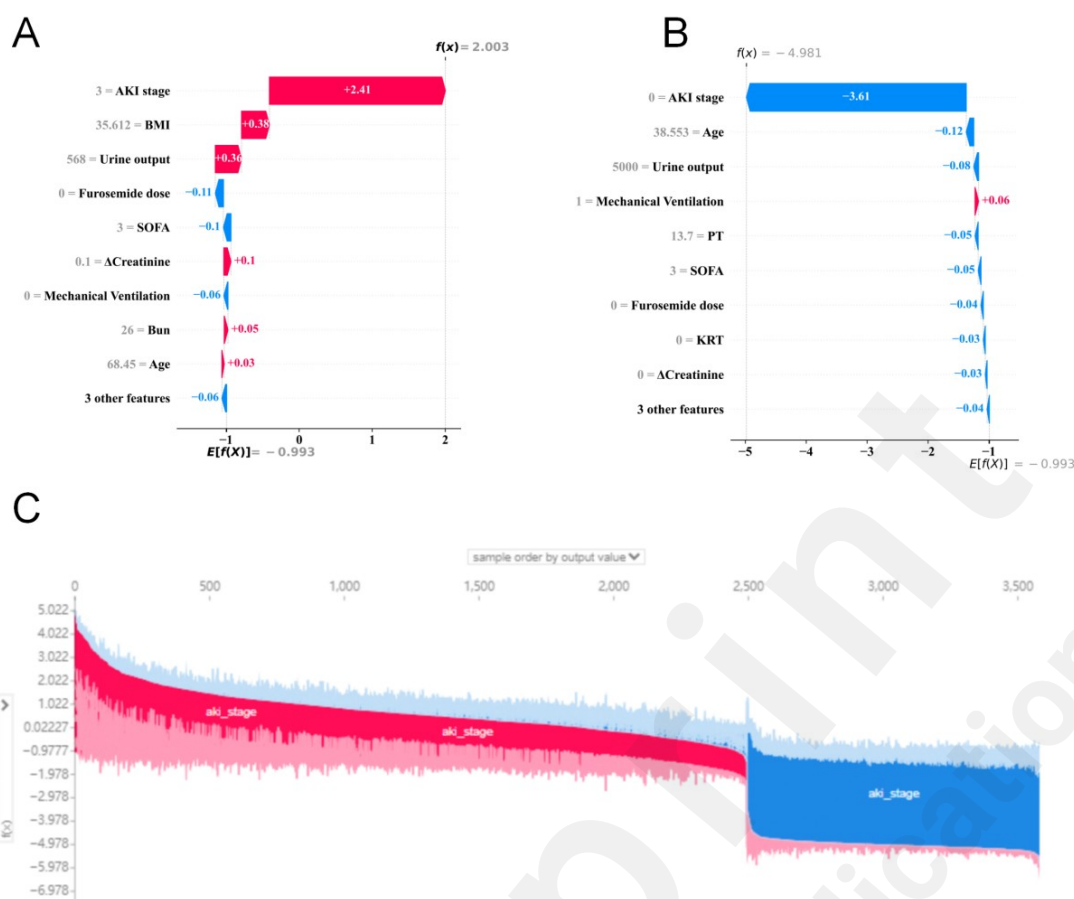


Figure 6: Local model explanation by the SHAP method. (A) represented the individual septic patients towards the “Persistent SA-AKI” class; (B) represented the individual septic patients towards the “non-Persistent SA-AKI” class. (C) Force plot for the internal validation set. Each patient was represented by the x-axis, while the feature contributions were represented by the y-axis: an increased red part for each individual patient represented a greater probability towards the decision of “Persistent SA-AKI”. SA-AKI: sepsis associated acute kidney injury; AKI: acute kidney injury; Δ Creatinine: Changes in creatinine within 24 hours; BMI: body mass Index; SOFA: sequential organ failure assessment; KRT: kidney renal therapy.

Convenient application for clinical utility

For the convenience of clinical application, the final prediction model has been implemented into a web-based application, as shown in Figure 7. When the actual values of the nine features required by the model are inputted, this app will automatically predict an individual patient's risk of persistent SA-AKI. In addition, it displays an interpretive force plot for each patient to indicate which features contribute to decisions about persistent SA-AKI: blue features on the right push predictions towards "non-persistent SA-AKI", while red features on the left push predictions towards "persistent SA-AKI". This web application can be accessed online [61] (The code for building the web version of the application is shown in Multimedia Appendix 9).

Persistent SA-AKI Prediction Application

Please enter the details for Persistent SA-AKI prediction

AKI Stage

3.0

-

+

BMI

27.0

-

+

Age

40

-

+

ΔCreatinine

2.00

-

+

SOFA Score

5

-

+

KRT

No

▼

Urine Output (mL)

500.0

-

+

Lactate (mmol/L)

1.5

-

+

Mechanical Ventilation

No

▼

Furosemide Dose (mg)

0.0

-

+

BUN (mg/dL)

13.0

-

+

PT (sec)

12.0

-

+

Predict

Prediction Result: Persistent SA-AKI

Prediction Probability: 84.59%

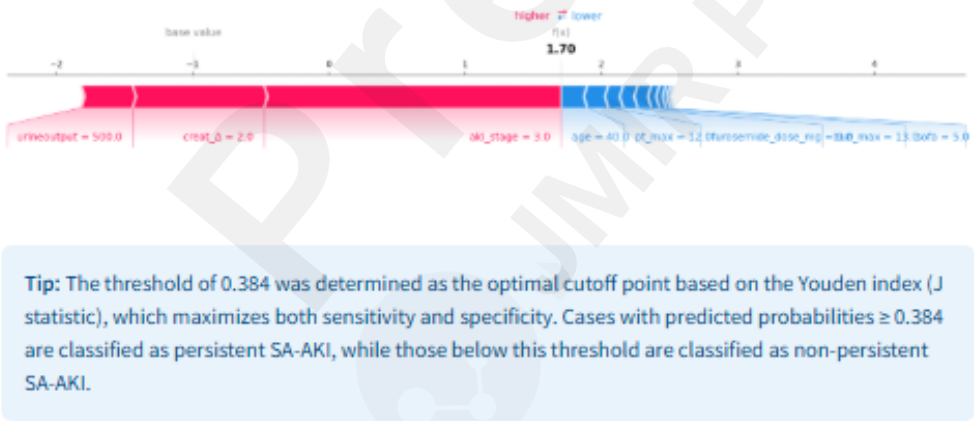


Figure 7: Convenient application for clinical utility. The convenient application of the final GBM model with 9 features is available for Persistent SA-AKI prediction. When entering the actual values of the 9 features, this application automatically displays the probability of 82.35%. Meanwhile, the force plot for individual child indicates the features that contribute to the decision of “Persistent SA-AKI”: the blue features on the right are the features pushing the prediction towards the “non-Persistent SA-AKI” class, while the red features on the left are pushing the prediction towards the “Persistent AKI” class. SA-AKI: sepsis-associated acute kidney injury.

Discussion

Principal Findings

This study not only established a model of persistent SA-AKI using a large database septic cohort, but also validated the model with retrospective and prospective cohorts. The model uses routine clinical data to predict persistent SA-AKI in septic patients in the ICU environment quite

well, and its diagnostic performance is even superior to that of CCL14, a biomarker known for diagnosing persistent SA-AKI.

Persistent SA-AKI has poor clinical prognosis, early identification and intervention are crucial. However, single-method diagnoses have limited effectiveness, including traditional urine indicators, renal ultrasound and biomarkers [10-19]. ML is an appealing method for predicting persistent SA-AKI. ML technology is a powerful computational method for handling complex and extensive data, as it can manage highly variable datasets while understanding complex relationships between variables in a flexible and trainable manner. EMR data are more easily accessible and accurate clinical data for clinicians and researchers. When combined with sophisticated ML algorithms, they can promote the development of clinical prediction models [40]. In this study, among eight ML models tested, the Gradient Boosting Machine (GBM) had the best Area Under Curve (AUC), demonstrating good net benefits along with higher feature reduction threshold probabilities. GBM consists of many simple decision trees, making it highly interpretable. Multiple studies have proven that GBM methods hold great predictive value within medical fields [41-43]. In our research, we used GBM to develop our final model featuring nine characteristics that can be easily obtained or evaluated.

The number of features included in any given model remains elusive due to a lack of guidelines or consensus on selecting features for predictive models. While more features may provide additional information for predictive models, including numerous ones could limit their clinical application; incorporating non-causal features might reduce prediction accuracy [44]. Using SHAP methodology assisted us in feature selection, resulting in our final simplified yet convenient ML predictive model, which can be easily applied to facilitate clinical decision-making for septic patients.

This study aims to apply ML algorithms to predict the occurrence of persistent AKI in septic patients. In our complex GBM model, we found that factors such as AKI stage, creatinine change value, urine output, diuretic dosage, mechanical ventilation, KRT, BMI, SOFA score, Bun, lactate, PT and age are associated with the development of persistent SA-AKI. Previous studies have shown a positive correlation between the severity of AKI in patients with this subtype and persistent AKI. Compared to transient AKI, patients with persistent AKI have higher stages of AKI and SOFA scores [45, 46]. The presence of a higher AKI stage and SOFA score is indicative of more severe kidney injury. As the severity of the injury escalates, the structural and functional impairment of the kidney is exacerbated, leading to an increase in the number of damaged renal units. The self-repairing and compensatory capacity of the kidney gradually diminishes, increasing the likelihood of developing persistent acute kidney injury. The higher stage is frequently accompanied by more severe inflammatory reactions and alterations in renal hemodynamics, providing a pathological foundation for the progression of persistent kidney injury. Our findings align with these previous studies; SHAP values from our GBM model suggest that higher stages of AKI and SOFA scores contribute to the occurrence of persistent SA-AKI. Urine output is considered one of the most important indicators for diagnosing AKI, which is consistent with KDIGO's recommendations. Oliguria is a significant risk factor for developing persistent SA-AKI, decreased urine output is indicative of inadequate renal perfusion or renal tubular dysfunction, which in turn results in decreased urine production. Inadequate renal perfusion results in oxygen deprivation of renal tissues, triggering cell damage and apoptosis. In contrast, renal tubular dysfunction affects the concentration and excretion of urine, thereby preventing the timely excretion of metabolites, which accumulate in the body and cause further damage to the kidneys, increasing the risk of persistent acute kidney injury. However, there remains controversy over whether using diuretics has any association with preventing or treating AKI [47-50]. Theoretically speaking, though, diuretics can prevent AKI by reducing glomerular filtration rate (GFR) and tubular reabsorption while also decreasing medullary oxygenation [51-53]. Some researchers believe that diuretics can act as renal vasodilators, thereby preventing AKI [54]. However, contrary to this belief, we found through our research that the use of diuretics plays an important role in influencing outcomes related to AKI. Consistent with our findings, several newly

published studies have discovered increased risks associated with usage of Diuretics leading towards AKI. A meta-analysis conducted by Ho & Sheridan involving 9 different studies showed that diuretics have no significant clinical benefits in preventing or treating AKI in adults. Worse still, treatment with diuretics may prolong hospital stays, and high doses of diuretics carry a short-term risk of deafness or tinnitus, which is harmful for ICU patients [55]. Therefore, renal function must be monitored during the course of treatment with diuretics. Creatinine change as a core indicator of RAI has been proven to predict the occurrence of AKI and persistent AKI in both adults and children [56-58]. The RAI is determined by subtle changes in the patient's condition and kidney function. It represents a special concept close to high-risk patients for persistent AKI. Our study also suggests that creatinine changes within 24 hours are associated with persistent SA-AKI. It's worth noting that mechanical ventilation is significantly associated with persistent SA-AKI, a significant increase in creatinine indicates a substantial impairment to kidney filtration function, potentially resulting from renal parenchyma damage and a decline in glomerular filtration rate. The inability of the kidneys to effectively remove creatinine leads to the accumulation of metabolic waste products within the body, which in turn can result in further damage to kidney cells, the triggering of an inflammatory response and oxidative stress, and the exacerbation of acute kidney injury. Positive pressure ventilation (PPV) is commonly used to provide oxygenation, ventilation, and airway protection support for critically ill patients. However, PPV has long been considered potentially harmful to the kidneys [59]. This could be due to three reasons: firstly, PPV might increase intrathoracic pressure, thereby reducing venous return, cardiac output and renal perfusion; secondly mechanical ventilation can induce release of certain neurohormones affecting renin-angiotensin system leading towards reduced renal blood flow and eGFR; thirdly any form of mechanical ventilation under any volume or pressure can trigger cascade inflammation including multiple leukocyte interleukins tumor necrosis factor-alpha and Fas ligand which might lead towards AKI. Aging and obesity are common risk factors for AKI. Obesity leads towards glomerular hypoperfusion increasing single nephron hemodynamics metabolic burden, activating adipocyte inflammation oxidative stress thus increasing risks related towards progression into AKI [7]. In sepsis, the inflammatory response activates the coagulation system, leading to extensive microthrombosis (e.g. disseminated intravascular coagulation, DIC) and depletion of coagulation factors and platelets, which ultimately manifests as prolonged PT. Prolonged PT is often an indicator of sepsis severity and poor prognosis and correlates with persistent SA-AKI and increased mortality. Persistent elevation of lactate is an indicator of poor prognosis. In sepsis, microcirculatory dysfunction leads to tissue hypoxia, anaerobic metabolism produces lactate, and hyperlactatemia further exacerbates acidosis and renal tubular injury, leading to the development of persistent acute kidney injury. With advancing age, the kidney undergoes a series of structural and functional changes, including a decrease in the number of renal units, glomerulosclerosis, tubular atrophy, and a decline in the reserve function and compensatory capacity of the kidney. Consequently, when subjected to identical injury factors, the kidneys of elderly individuals are more susceptible to acute kidney injury. Due to their diminished repair capacity, it is more challenging to recover, and there is an increased risk of persistent acute kidney injury. Since these indicators are easy to assess at the time of admission, they can serve as convenient predictive factors for critically ill septic patients developing persistent SA-AKI. In summary, the majority of the model's key features, including AKI classification, creatinine change values, urine output, mechanical ventilation, KRT, age, BMI, Bun, lactate, PT and SOFA scores, are consistent with clinical risk and contribute to the worsening of AKI. However, our study also identified some discrepancies that are of concern. For instance, the established correlation between diuretic dosage and AKI recovery deviates from conventional wisdom. Higher doses of diuretics did not promote recovery of renal function as expected and were even associated with a poorer prognosis to some extent. This finding calls into question the conventional wisdom surrounding the effectiveness of diuretics in the treatment of AKI and suggests that clinicians should exercise greater caution when prescribing diuretics.

Comparison With Prior Work

As previously reported, some studies have established AKI prediction models based on machine learning (ML) methods [21–26], which have focused on untyped acute kidney injury (AKI). This status quo is not conducive to guiding the clinic towards personalized treatments. Only a few studies have employed machine learning models to predict persistent AKI[27, 60]. For instance, Ding et al. [60] found that serum albumin, chronic kidney disease, AKI stage, SOFA score, lactate on the first day, and kidney replacement therapy were significantly associated with persistent AKI through LASSO regression and SVM-RFE analyses using a free database. Utilizing these predictors, a column-line graph was constructed. The diagnostic efficacy of this column chart in predicting persistent AKI in the training set was 0.730, whereas in the external validation cohort it was only 0.702. It is noteworthy that several of the factors examined in this study, including AKI staging, SOFA score, lactate and kidney replacement therapy, bear a resemblance to those observed in our own study. Nevertheless, the predictive model developed in this study demonstrated unsatisfactory diagnostic efficacy in both the training and validation cohorts. Jiang et al. [27] conducted a retrospective analysis of 955 patients with post-surgical co-morbid AKIs, and predicted the occurrence of persistent AKIs using a model integrated with three machine learning methods. The model demonstrated a diagnostic value of 0.86 in the training set, yet a comparatively modest diagnostic value of 0.693 in an external validation cohort. This limitation directly restricts the generalizability of the model. In comparison with preceding studies, the model that was constructed has a number of advantages. Firstly, the model demonstrates remarkable stability, with a diagnostic value for predicting persistent AKI above 0.85 in both the training and validation cohorts, thus demonstrating excellent diagnostic performance. Secondly, we seek to address the 'black box' problem of machine learning (ML) methods that are difficult to interpret directly by using the Shapley Additive explanations (SHAP) method to profile machine learning models. This method can be interpreted at two levels: the global interpretation and the local interpretation. The model can be interpreted at both the global and local levels: the global interpretation describes the overall functionality of the model, while the local interpretation makes predictions specific to individual patients with sepsis through the input of personalized data. This facilitates a more profound comprehension of the machine learning model's inner workings by clinicians, thereby enhancing their propensity to make informed medical decisions. Thirdly, the prediction models constructed in previous studies are difficult to use in clinical applications. In response to this, an online APP was developed based on the Flow Light framework to facilitate clinicians' access and assist them in clinical decision-making. Fourthly, there is a paucity of studies that have hitherto compared self-developed models with established markers in prospective cohorts. In this study, a machine learning model was compared with a marker for CCL14, a chemokine belonging to the CC chemokine family that plays an important role in immune regulation and inflammatory responses [19]. It has been established that CCL14 is closely associated with the degree of renal tubular injury and disease progression, and its elevated urinary concentration may reflect the activation of local inflammatory responses in renal tissues, which are involved in the process of renal tubular epithelial cell apoptosis and interstitial fibrosis. The available studies have demonstrated that CCL14 exhibits good diagnostic efficacy for persistent severe acute kidney injury (SA-AKI), with the efficacy of urinary CCL14 alone for the prediction of persistent SA-AKI being above 0.80[19, 61], a figure that is currently recognized as the most effective predictor of persistent severe AKI[20, 62]. The diagnostic efficacy of the model can be reflected by comparing it with CCL14. In comparison, our machine learning model outperforms CCL14, further highlighting the superiority and reliability of our model.

In summary, we have successfully established an interpretable ML model that predicts persistent SA-AKI in septic patients in the ICU environment based on easily extractable clinical data. The final GBM has good predictive ability in both internal and external validation, and its diagnostic efficacy is even superior to CCL14 in prospective cohort studies. Further randomized and controlled studies

are needed to determine whether individualized and timely treatment measures based on the final prediction model can improve the prognosis of patients with persistent SA-AKI.

Limitations

Of course, we must acknowledge some limitations in this study. Firstly, as a common issue among hospitalized patients, the baseline Scr values for most patients are unobtainable. We used an estimation equation to determine the patient's baseline Scr, a method that has been employed in many AKI studies to minimize errors as much as possible. Secondly, we only included information on demographics, vital signs, and laboratory values, which limited the performance of the model and prevented it from discovering more potential factors that contribute to persistent SA-AKI. Lastly, our model only includes septic patients; however, AKI is a syndrome with various causes of onset, so it remains unclear whether this model can predict persistent AKI caused by other factors - further verification is needed.

Conclusions

In summary, we have successfully established an interpretable GBM model for predicting persistent SA-AKI based on readily extractable clinical data. This model demonstrated good predictive ability in both internal and external validation cohorts and outperformed the biomarker CCL14 in prospective cohort validation.

Author's contributions

All the authors participated in literature retrieval and viewpoint discussion in this article. Wei Jiang was responsible for the research design, paper writing, data extraction, and data analysis of the study as the main contributor, Yao Sheng Zhang was responsible for the extraction of the public database, Jiayi Weng and Siqi Liu were responsible for the machine learning analyses, Lin Song, Xianghui Li, Siqi Xu, Quan Yuan, Jing Wang, Keran Shi, Luanluan Li, Chuanqing Zhang was responsible for the collection of hospital data, Yongwei Zhang and Quan Yuan organized the charts, Jun Shao, Jiangquan Yu and Ruiqiang Zheng participated in the discussion of the topic, and the research design. Ruiqiang Zheng is responsible for all the study work as the guarantor.

Funding

This work was supported by Key Projects of Jiangsu Province Traditional Chinese Medicine Science and Technology Development Plan (ZD202427); Jiangsu Provincial Medical Key Discipline Cultivation Unit (JSDW20221); Yangzhou Commission of Health key research project (20231-02); Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX24_2330, SJCX24_2348); The Development Fund of the Affiliated Hospital of Xuzhou Medical University (XYFM202401);

Availability of data and materials

Data are available upon reasonable request.

Ethics approval and consent to participate

Given that this study was based on an analysis of a third-party anonymous public database and prior approval was obtained from the institutional review committee, an ethical review was therefore not required. The data collected by the ICU at Northern Jiangsu People's Hospital has been approved by its Ethics Committee (ID number: 2021ky230) and registered on the China Clinical Trial Registration Platform (Registration number: ChiCTR2100050540).

Consent for publication

My co-authors have all contributed to approve of this submission.

620 Competing interests

The authors have no competing interest to declare.

Multimedia Appendix 1

625 Demographic characteristics, vital sign measurements, and laboratory variables with the percentage of missing.

Multimedia Appendix 2

Multimedia Appendix 2 Heat map of Spearman's correlation analyses among variables Variables of sepsis patients from the derivation cohort were included in Spearman's correlation analyses.

Multimedia Appendix 3

630 Comparison of demographic and clinical characteristics among the training, internal validation, and external validation cohorts.

Multimedia Appendix 4

Performance of the ML models for persistent SA-AKI prediction.

Multimedia Appendix 5

635 SHAP summary dot plots of the 12 features of the top four best-performing ML models.

Multimedia Appendix 6

Performance of the GBM model with varied numbers of features for Persistent SA-AKI prediction.

Multimedia Appendix 7

Predictive performance of the GBM model after reducing features.

640 Multimedia Appendix 8

Predictive performance of the final GBM model in the cross validation.

Multimedia Appendix 9

Web app codes

645 Reference

- [1] PEERAPORNATANA S, MANRIQUE-CABALLERO C L, GÓMEZ H, et al. Acute kidney injury from sepsis: current concepts, epidemiology, pathophysiology, prevention and treatment [J]. *Kidney Int*, 2019, 96(5): 1083-99.
- [2] ZARBOCK A, NADIM M K, PICKKERS P, et al. Sepsis-associated acute kidney injury: consensus report of the 28th Acute Disease Quality Initiative workgroup [J]. *Nat Rev Nephrol*, 2023, 19(6): 401-17.
- [3] PETERS E, ANTONELLI M, WITTEBOLE X, et al. A worldwide multicentre evaluation of the influence of deterioration or improvement of acute kidney injury on clinical outcome in critically ill patients with and without sepsis at ICU admission: results from The Intensive Care Over Nations audit [J]. *Crit Care*, 2018, 22(1): 188.
- [4] BELLOMO R, KELLUM J A, RONCO C, et al. Acute kidney injury in sepsis [J]. *Intensive Care Med*, 2017, 43(6): 816-28.
- [5] POSTON J T, KOYNER J L. Sepsis associated acute kidney injury [J]. *Bmj*, 2019, 364:

k4891.

- 660 [6] CHAWLA L S, BELLOMO R, BIHORAC A, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup [J]. *Nat Rev Nephrol*, 2017, 13(4): 241-57.
- [7] UHEL F, PETERS-SENGERS H, FALAHI F, et al. Mortality and host response aberrations associated with transient and persistent acute kidney injury in critically ill patients with sepsis: a prospective cohort study [J]. *Intensive Care Med*, 2020, 46(8): 1576-89.
- 665 [8] KELLUM J A, SILEANU F E, BIHORAC A, et al. Recovery after Acute Kidney Injury [J]. *Am J Respir Crit Care Med*, 2017, 195(6): 784-91.
- [9] CHAWLA L S, EGGERS P W, STAR R A, et al. Acute kidney injury and chronic kidney disease as interconnected syndromes [J]. *N Engl J Med*, 2014, 371(1): 58-66.
- 670 [10] FORNI L G, JOANNIDIS M. IDEAL timing of renal replacement therapy in critical care [J]. *Nat Rev Nephrol*, 2019, 15(1): 5-6.
- [11] DARMON M, VINCENT F, DELLAMONICA J, et al. Diagnostic performance of fractional excretion of urea in the evaluation of critically ill patients with acute kidney injury: a multicenter cohort study [J]. *Crit Care*, 2011, 15(4): R178.
- 675 [12] PONS B, LAUTRETTE A, OZIEL J, et al. Diagnostic accuracy of early urinary index changes in differentiating transient from persistent acute kidney injury in critically ill patients: multicenter cohort study [J]. *Crit Care*, 2013, 17(2): R56.
- [13] VANMASSENHOVE J, GLORIEUX G, HOSTE E, et al. Urinary output and fractional excretion of sodium and urea as indicators of transient versus intrinsic acute kidney injury during early sepsis [J]. *Crit Care*, 2013, 17(5): R234.
- 680 [14] DARMON M, BOURMAUD A, REYNAUD M, et al. Performance of Doppler-based resistive index and semi-quantitative renal perfusion in predicting persistent AKI: results of a prospective multicenter study [J]. *Intensive Care Med*, 2018, 44(11): 1904-13.
- [15] DEWITTE A, JOANNÈS-BOYAU O, SIDOBRE C, et al. Kinetic eGFR and Novel AKI Biomarkers to Predict Renal Recovery [J]. *Clin J Am Soc Nephrol*, 2015, 10(11): 1900-10.
- 685 [16] TITECA-BEAUPORT D, DAUBIN D, VAN VONG L, et al. Urine cell cycle arrest biomarkers distinguish poorly between transient and persistent AKI in early septic shock: a prospective, multicenter study [J]. *Crit Care*, 2020, 24(1): 280.
- [17] GARNIER F, DAUBIN D, LARCHER R, et al. Reversibility of Acute Kidney Injury in Medical ICU Patients: Predictability Performance of Urinary Tissue Inhibitor of Metalloproteinase-2 x Insulin-Like Growth Factor-Binding Protein 7 and Renal Resistive Index [J]. *Crit Care Med*, 2020, 48(4): e277-e84.
- 690 [18] JIANG W, YU J, WANG H, et al. The value of urine cell cycle arrest biomarkers to predict persistent acute kidney injury: A systematic review and meta-analysis [J]. *Clin Nephrol*, 2021, 96(6): 327-36.
- 695 [19] HOSTE E, BIHORAC A, AL-KHAFAJI A, et al. Identification and validation of biomarkers of persistent acute kidney injury: the RUBY study [J]. *Intensive Care Med*, 2020, 46(5): 943-53.
- [20] JIANG W, LIAO T, YU J, et al. Predictability performance of urinary C-C motif chemokine ligand 14 and renal resistive index for persistent sepsis-associated acute kidney injury in ICU patients [J]. *Int Urol Nephrol*, 2023, 55(8): 1995-2003.
- 700 [21] GAO W, WANG J, ZHOU L, et al. Prediction of acute kidney injury in ICU with gradient boosting decision tree algorithms [J]. *Comput Biol Med*, 2022, 140: 105097.
- [22] KOYNER J L, CAREY K A, EDELSON D P, et al. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model [J]. *Crit Care Med*, 2018, 46(7): 1070-7.
- 705 [23] ALFIERI F, ANCONA A, TRIPEPI G, et al. Continuous and early prediction of future moderate and severe Acute Kidney Injury in critically ill patients: Development and multi-centric, multi-national external validation of a machine-learning model [J]. *PLoS One*, 2023,

18(7): e0287398.

- 710 [24] GONG K, LEE H K, YU K, et al. A prediction and interpretation framework of acute kidney injury in critical care [J]. *J Biomed Inform*, 2021, 113: 103653.
- [25] SHAWWA K, GHOSH E, LANIUS S, et al. Predicting acute kidney injury in critically ill patients using comorbid conditions utilizing machine learning [J]. *Clin Kidney J*, 2021, 14(5): 1428-35.
- 715 [26] FLECHET M, GüIZA F, SCHETZ M, et al. AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated lipocalin [J]. *Intensive Care Med*, 2017, 43(6): 764-73.
- [27] JIANG X, HU Y, GUO S, et al. Prediction of persistent acute kidney injury in postoperative intensive care unit patients using integrated machine learning: a retrospective cohort study [J]. *Sci Rep*, 2022, 12(1): 17134.
- 720 [28] JOHNSON A E W, BULGARELLI L, SHEN L, et al. MIMIC-IV, a freely accessible electronic health record dataset [J]. *Sci Data*, 2023, 10(1): 1.
- [29] JOHNSON A E, POLLARD T J, SHEN L, et al. MIMIC-III, a freely accessible critical care database [J]. *Sci Data*, 2016, 3: 160035.
- 725 [30] POLLARD T J, JOHNSON A E W, RAFFA J D, et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research [J]. *Sci Data*, 2018, 5: 180178.
- [31] SINGER M, DEUTSCHMAN C S, SEYMOUR C W, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) [J]. *Jama*, 2016, 315(8): 801-10.
- 730 [32] RATH E P, UDY A A, BAILEY M, et al. Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit [J]. *Jama*, 2017, 317(3): 290-300.
- 735 [33] LAMBDEN S, LATERRE P F, LEVY M M, et al. The SOFA score-development, utility and challenges of accurate assessment in clinical trials [J]. *Crit Care*, 2019, 23(1): 374.
- [34] SEYMOUR C W, LIU V X, IWASHYNA T J, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) [J]. *Jama*, 2016, 315(8): 762-74.
- 740 [35] KELLUM J A, LAMEIRE N. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1) [J]. *Crit Care*, 2013, 17(1): 204.
- [36] SIEW E D, IKIZLER T A, MATHENY M E, et al. Estimating baseline kidney function in hospitalized patients with impaired kidney function [J]. *Clin J Am Soc Nephrol*, 2012, 7(5): 712-9.
- 745 [37] HUBER M, OZRAZGAT-BASLANTI T, THOTTAKKARA P, et al. Cardiovascular-Specific Mortality and Kidney Disease in Patients Undergoing Vascular Surgery [J]. *JAMA Surg*, 2016, 151(5): 441-50.
- [38] ANGELI P, GINÈS P, WONG F, et al. Diagnosis and management of acute kidney injury in patients with cirrhosis: revised consensus recommendations of the International Club of Ascites [J]. *J Hepatol*, 2015, 62(4): 968-74.
- 750 [39] ZHOU D, JIANG Y, ZHONG X, et al. A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis [J]. *Nat Genet*, 2020, 52(11): 1239-46.
- [40] GOECKS J, JALILI V, HEISER L M, et al. How Machine Learning Will Transform Biomedicine [J]. *Cell*, 2020, 181(1): 92-101.
- 755 [41] WANG S, ZHANG T, LI Z, et al. Exploring pollutant joint effects in disease through interpretable machine learning [J]. *J Hazard Mater*, 2024, 467: 133707.
- [42] NIGO M, RASMY L, MAO B, et al. Deep learning model for personalized prediction of

positive MRSA culture using time-series electronic health records [J]. *Nat Commun*, 2024, 15(1): 2036.

[43] DEVANARAYAN V, YE Y, CHARIL A, et al. Predicting clinical progression trajectories of early Alzheimer's disease patients [J]. *Alzheimers Dement*, 2024, 20(3): 1725-38.

[44] LI Y, SPERRIN M, ASHCROFT D M, et al. Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar [J]. *Bmj*, 2020, 371: m3919.

[45] JIANG W, ZHANG C, YU J, et al. Development and validation of a nomogram for predicting in-hospital mortality of elderly patients with persistent sepsis-associated acute kidney injury in intensive care units: a retrospective cohort study using the MIMIC-IV database [J]. *BMJ Open*, 2023, 13(3): e069824.

[46] WANG M, WANG X, ZHU B, et al. The effects of timing onset and progression of AKI on the clinical outcomes in AKI patients with sepsis: a prospective multicenter cohort study [J]. *Ren Fail*, 2023, 45(1): 1-10.

[47] ZHAO G J, XU C, YING J C, et al. Association between furosemide administration and outcomes in critically ill patients with acute kidney injury [J]. *Crit Care*, 2020, 24(1): 75.

[48] SHEN Y, WU M. Loop diuretic use in patients with AKI: different severity, different response [J]. *Crit Care*, 2018, 22(1): 202.

[49] GAO S, YANG L, WANG Z. Do critically ill patients with AKI benefit from furosemide? Further real-world evidence from a large multi-center database [J]. *Crit Care*, 2020, 24(1): 253.

[50] JOANNIDIS M, KLEIN S J, OSTERMANN M. 10 myths about frusemide [J]. *Intensive Care Med*, 2019, 45(4): 545-8.

[51] BREZIS M, ROSEN S. Hypoxia of the renal medulla--its implications for disease [J]. *N Engl J Med*, 1995, 332(10): 647-55.

[52] EPSTEIN F H, PRASAD P. Effects of furosemide on medullary oxygenation in younger and older subjects [J]. *Kidney Int*, 2000, 57(5): 2080-3.

[53] PEIXOTO A J. Update in Nephrology and Hypertension: Evidence Published in 2015 [J]. *Ann Intern Med*, 2016, 164(9): W42-7.

[54] KELLUM J A. Diuretics in acute renal failure: protective or deleterious [J]. *Blood Purif*, 1997, 15(4-6): 319-22.

[55] HO K M, SHERIDAN D J. Meta-analysis of frusemide to prevent or treat acute renal failure [J]. *Bmj*, 2006, 333(7565): 420.

[56] MATSUURA R, SRISAWAT N, CLAURE-DEL GRANADO R, et al. Use of the Renal Angina Index in Determining Acute Kidney Injury [J]. *Kidney Int Rep*, 2018, 3(3): 677-83.

[57] CRUZ D N, FERRER-NADAL A, PICCINNI P, et al. Utilization of small changes in serum creatinine with clinical risk factors to assess the risk of AKI in critically ill adults [J]. *Clin J Am Soc Nephrol*, 2014, 9(4): 663-72.

[58] BASU R K, ZAPPITELLI M, BRUNNER L, et al. Derivation and validation of the renal angina index to improve the prediction of acute kidney injury in critically ill children [J]. *Kidney Int*, 2014, 85(3): 659-67.

[59] ZACCHETTI L, BRIVIO M, MEZZAPESA M, et al. The effect of positive pressure ventilation on Acute Kidney Injury in COVID-19 patients with Acute Respiratory Distress Syndrome. An observational study [J]. *Blood Purif*, 2024.

[60] DING C, HU T. Development and External Verification of a Nomogram for Patients with Persistent Acute Kidney Injury in the Intensive Care Unit [J]. *Int J Gen Med*, 2021, 14: 5005-15.

[61] BAGSHAW S M, AL-KHAFAJI A, ARTIGAS A, et al. External validation of urinary C-C motif chemokine ligand 14 (CCL14) for prediction of persistent acute kidney injury [J]. *Crit Care*, 2021, 25(1): 185.

- 810 [62] KELLUM J A, BAGSHAW S M, DEMIRJIAN S, et al. CCL14 testing to guide clinical practice in patients with AKI: Results from an international expert panel [J]. J Crit Care, 2024, 82: 154816.
- [63] <https://persistent-sa-aki-prediction-model.streamlit.app/>

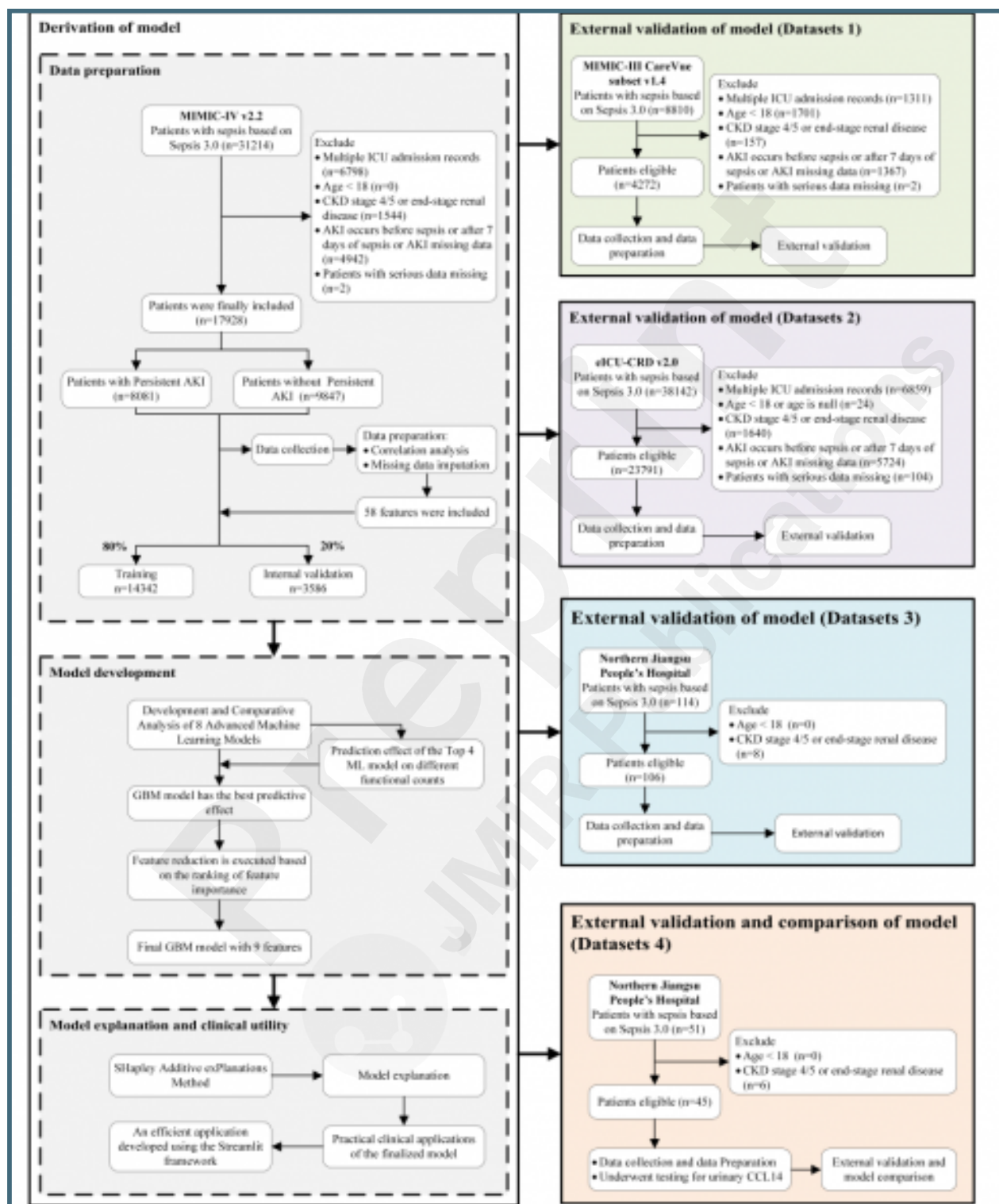
Abbreviations

- 815 AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
SA-AKI: sepsis associated acute kidney injury
KDIGO: Kidney Disease: Improving Global Outcomes
SCr: serum creatinine
- 820 SHAP: Shapley additive explanations
ML: machine learning
GBM: gradient boosting machine
CatBoost: categorical boosting
XGboost: extreme gradient boost
- 825 SVM: support vector machine
ANN: artificial neural network
DT: decision tree; KNN: K-nearest neighbor

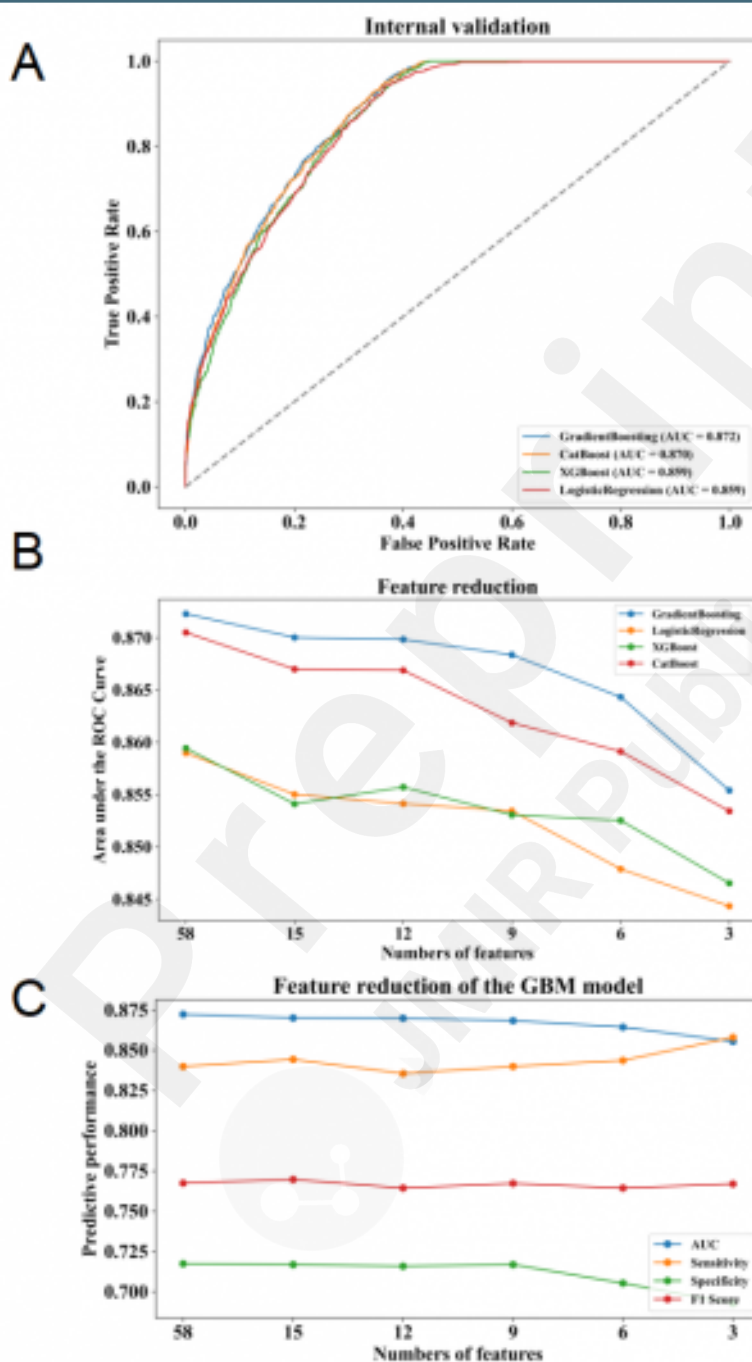
Supplementary Files

Figures

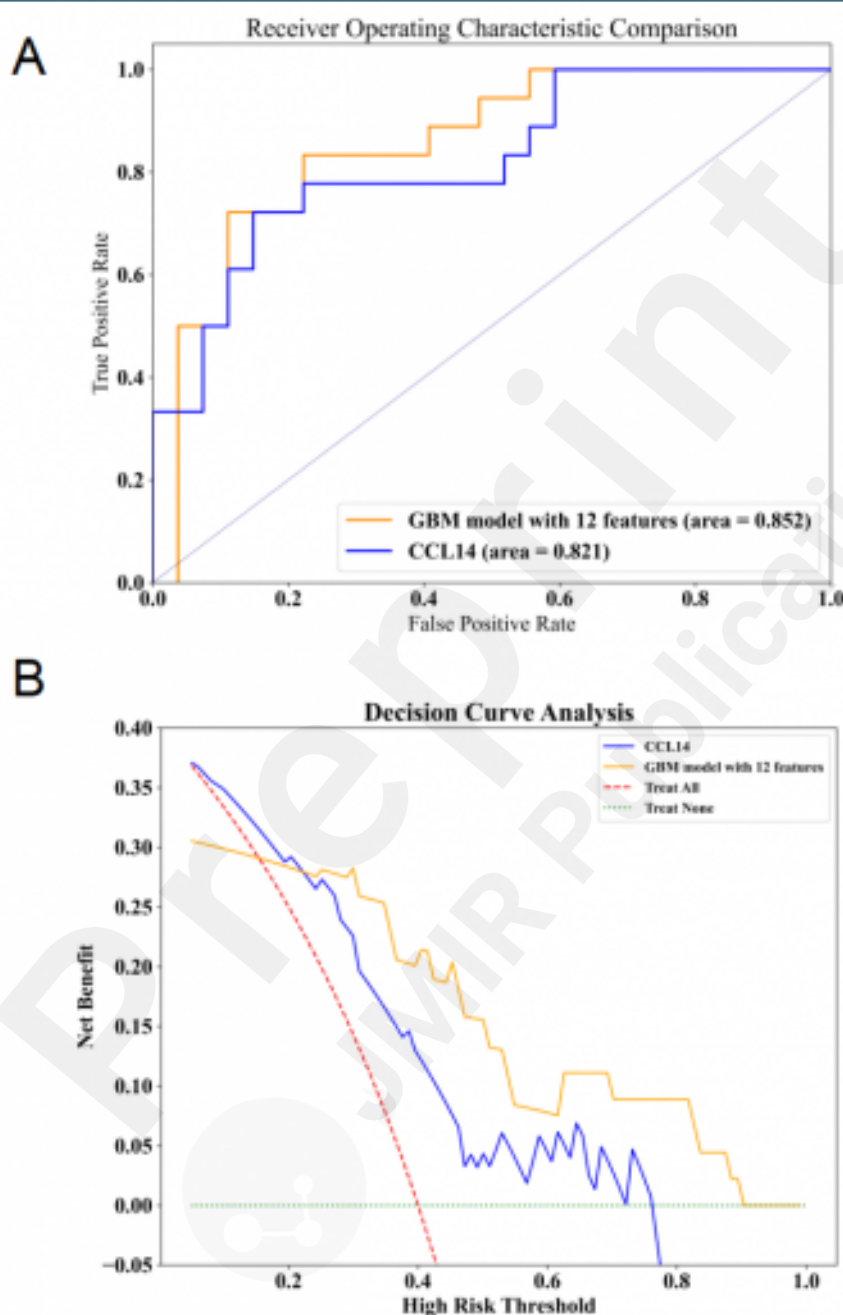
Study flow chart. Sepsis-associated acute kidney injury; CKD: chronic kidney disease; GBM: Gradient Boosting Machine; ICU: intensive care unit; SHAP: SHapley Additive explanation; CCL14: C-C motif chemokine ligand 14.



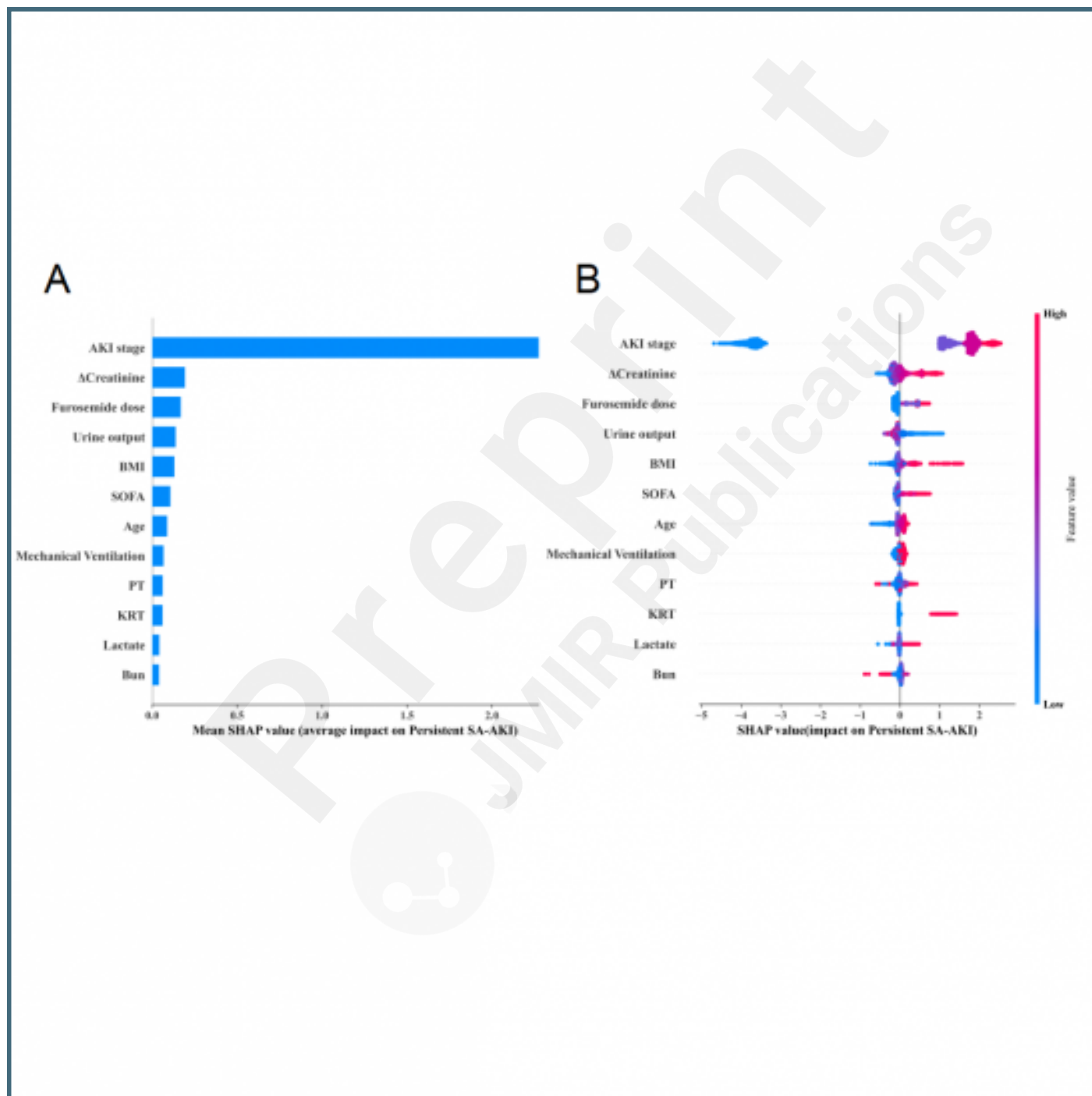
Performance of ML models to predict Persistent SA-AKI. (A) ROC curves of the top four best-performing ML models. (B) AUCs of the top four best performing ML models with varied numbers of features. (C) AUC, sensitivity, specificity, and F1 score of the GBM model with varied numbers of features. SA-AKI:sepsis associated acute kidney injury;AUC:area under the ROC curve;ROC: receiver-operating-characteristic;GBM: gradient boosting machine; XGboost: eXtreme gradient boosting; CatBoost: categorical boosting;ML: machine learning.



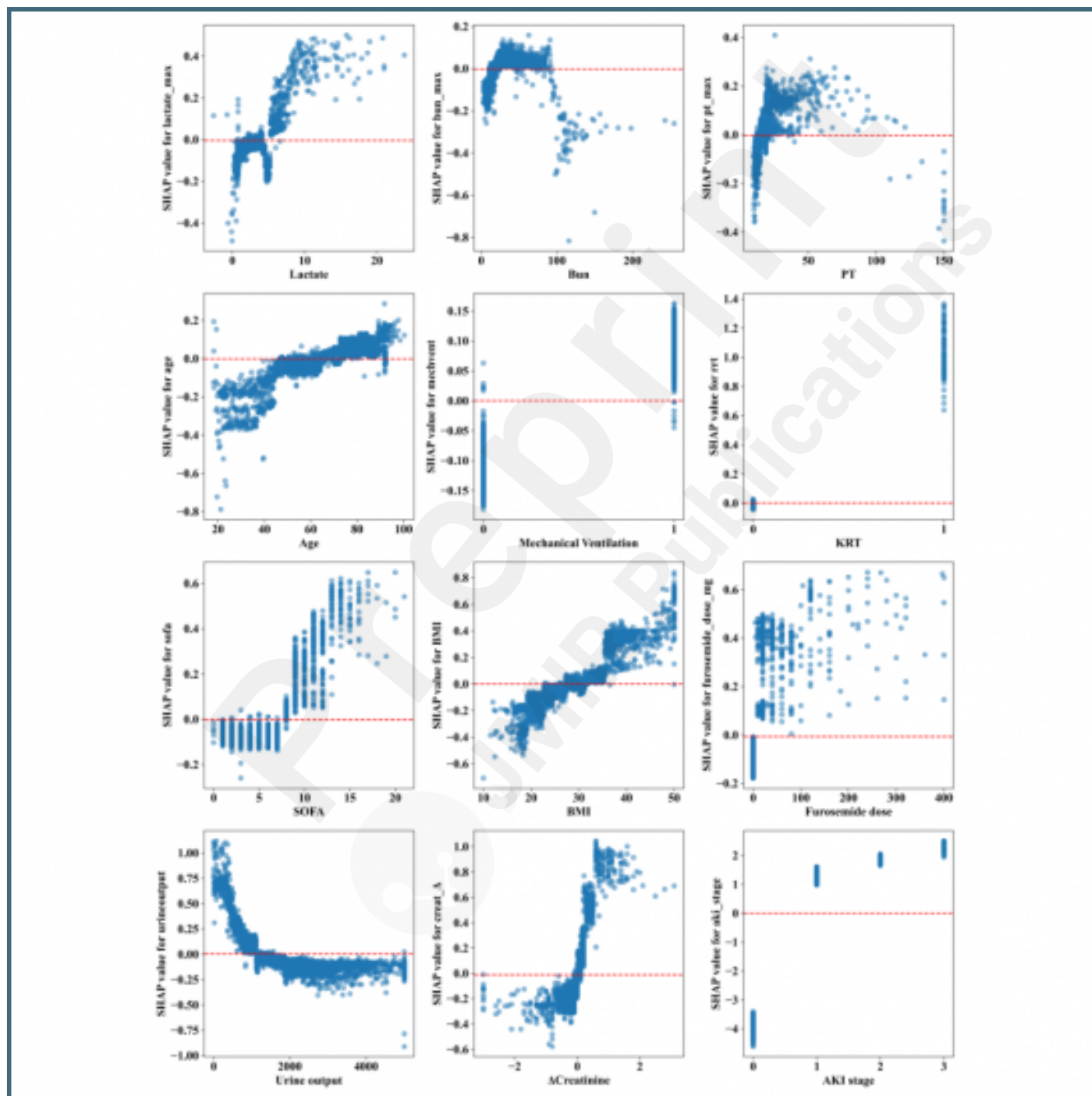
Comparison between final GBM model and CCL14. (A-B) ROC curves (A) and DCA curves (B) of CCL14 and the final GBM model with 8 features. These plots represented the predictive performance in the prospective external validation cohort. AUC: area under the ROC curve; DCA: decision curve analysis; ROC: receiver-operating-characteristic; GBM: gradient boosting machine; CCL14: C-C motif chemokine ligand 14;



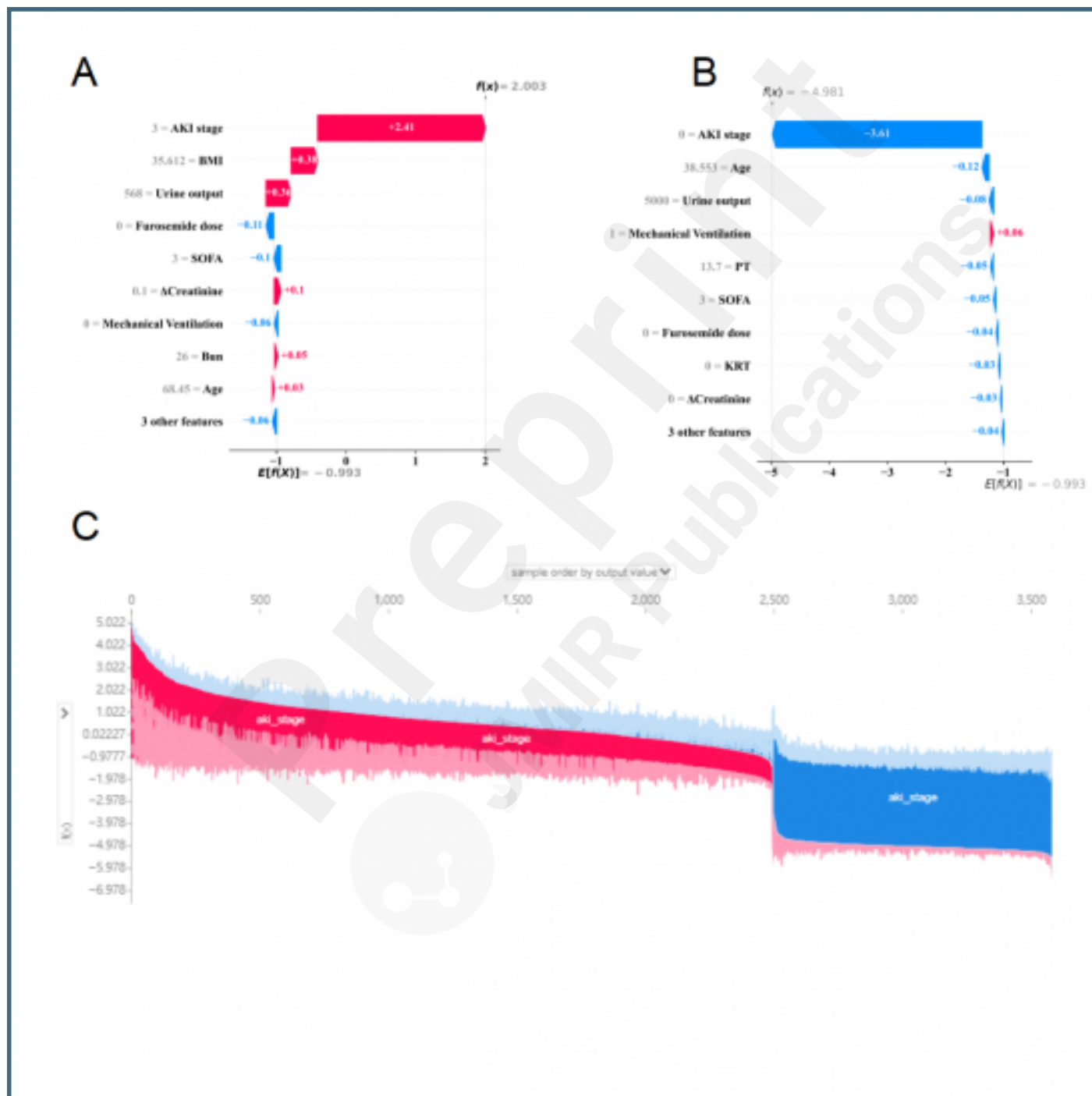
Global model explanation by the SHAP method. (A) SHAP summary bar plot. (B) SHAP summary dot plot. The probability of Persistent SA-AKI development increases with the SHAP value of a feature. A dot is made for SHAP value in the model for each single patient, so each patient has one dot on the line for each feature. The colors of the dots demonstrate the actual values of the features for each patient, as red means a higher feature value and blue means a lower feature value. The dots are stacked vertically to show density. SA-AKI:sepsis associated acute kidney injury;AKI:acute kidney injury; ?Creatinine:Changes in creatinine within 24 hours;BMI: body mass Index; SOFA: sequential organ failure assessment; KRT:kidney renal therapy;PT:Prothrombin time.



SHAP dependence plot. Each dependence plot shows how a single feature affects the output of the prediction model, and each dot represents a single patient. For example, SOFA scores above eight points or urine outputs less than 1380ml/24h or BMI indices above 24.92 or age older than 66 years push the decision towards the “Persistent SA-AKI” class. SHAP values are represented by the y-axis, and actual values are represented by the x-axis. The SHAP values for specific features exceeding zero push the decision towards the “Persistent SA-AKI” class. SA-AKI: sepsis associated acute kidney injury; AKI: acute kidney injury; ?Creatinine: Changes in creatinine within 24 hours; BMI: body mass Index; SOFA: sequential organ failure assessment; KRT: kidney renal therapy; PT: Prothrombin time.



Local model explanation by the SHAP method. (A) represented the individual septic patients towards the “Persistent SA-AKI” class, (B) represented the individual septic patients towards the “non- Persistent SA-AKI” class. (C) Force plot for the internal validation set. Each patient was represented by the x-axis, while the features’ contributions were represented by the y-axis: an increased red part for each individual patient represented a greater probability towards the decision of “Persistent SA-AKI”. SA-AKI: sepsis associated acute kidney injury; AKI: acute kidney injury; Δ Creatinine: Changes in creatinine within 24 hours; BMI: body mass Index; SOFA: sequential organ failure assessment; KRT: kidney renal therapy; PT: Prothrombin time.



Convenient application for clinical utility. The convenient application of the final GBM model with 12 features is available for Persistent SA-AKI prediction. When entering the actual values of the 12 features, this application automatically displays the probability of 82.35%. Meanwhile, the force plot for individual child indicates the features that contribute to the decision of “Persistent SA-AKI”: the blue features on the right are the features pushing the prediction towards the “non-Persistent SA-AKI” class, while the red features on the left are pushing the prediction towards the “Persistent AKI” class. SA-AKI: sepsis-associated acute kidney injury.

Persistent SA-AKI Prediction Application

Please enter the details for Persistent SA-AKI prediction

AKI Stage 3.0	BMI 27.0	Age 40
ΔCreatinine 2.00	SOFA Score 5	KRT No
Urine Output (mL) 500.0	Lactate (mmol/L) 1.5	Mechanical Ventilation No
Furosemide Dose (mg) 0.0	BUN (mg/dL) 13.0	PT (sec) 12.0

Predict

Prediction Result: Persistent SA-AKI

Prediction Probability: 84.59%

Force plot showing the contribution of 12 features to the prediction. The x-axis represents the force, with a baseline at 0. Features to the left (red) push towards Persistent SA-AKI, and features to the right (blue) push towards non-Persistent SA-AKI. The threshold is 0.384.

Tip: The threshold of 0.384 was determined as the optimal cutoff point based on the Youden index (J statistic), which maximizes both sensitivity and specificity. Cases with predicted probabilities ≥ 0.384 are classified as persistent SA-AKI, while those below this threshold are classified as non-persistent SA-AKI.

Multimedia Appendixes

Demographic characteristics, vital sign measurements, and laboratory variables with the percentage of missing.

URL: <http://asset.jmir.pub/assets/cc82c4da40d3328abcd8c955794946ac.docx>

Heat map of Spearman's correlation analyses among variables Variables of sepsis patients from the derivation cohort were included in Spearman's correlation analyses.

URL: <http://asset.jmir.pub/assets/87a60a30f63b065631e04547d99ca435.docx>

Comparison of demographic and clinical characteristics among the training, internal validation, and external validation cohorts.

URL: <http://asset.jmir.pub/assets/71d588116b837c0f6cdf2c8f5c998602.docx>

Performance of the ML models for persistent SA-AKI prediction.

URL: <http://asset.jmir.pub/assets/e47b503b14c9b331e3063bfbe19364b7.docx>

SHAP summary dot plots of the 12 features of the top four best-performing ML models.

URL: <http://asset.jmir.pub/assets/1a13b6ce0d9ee5a3fb27b6493a242f3f.docx>

Performance of the GBM model with varied numbers of features for Persistent SA-AKI prediction.

URL: <http://asset.jmir.pub/assets/1e6a42bbc30f7109a1d3b7c8214354ed.docx>

Predictive performance of the GBM model after reducing features.

URL: <http://asset.jmir.pub/assets/fbe76f67597428c9db30790f2ef414a3.docx>

Predictive performance of the final GBM model in the cross validation.

URL: <http://asset.jmir.pub/assets/974e58ea174511ff9022e628914e0dff.docx>

Web app codes.

URL: <http://asset.jmir.pub/assets/2eecf25406d9a1705691a95d176d2f0a.docx>