Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and external validation of an acute kidney injury risk model in older ICU patients with aspiration pneumonia

Hongbo Li
The Second Nanning People's Hospital
Qimeng Xu
The People's Hospital of Guangxi Zhuang Autonomous Region
Yilan Liang
The People's Hospital of Guangxi Zhuang Autonomous Region
Sitao Tan
The People's Hospital of Guangxi Zhuang Autonomous Region
Ying He
The People's Hospital of Guangxi Zhuang Autonomous Region
Junliang Qin
The People's Hospital of Guangxi Zhuang Autonomous Region
Shulin Xiang
The People's Hospital of Guangxi Zhuang Autonomous Region
Xiaoxia Liu
p2412574@mpu.edu.mo

The People's Hospital of Guangxi Zhuang Autonomous Region

Article

Keywords: aspiration pneumonia, older patients, acute kidney injury, intensive care unit, machine learning, external validation

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11140880/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Acute kidney injury (AKI) is a major complication in older patients with aspiration pneumonia. This retrospective multicohort study developed and externally validated an interpretable AKI risk assessment model for patients aged ≥ 65 years admitted to intensive care units (ICUs). The study included 3,125 patients from MIMIC-IV for model development and internal validation, and 1,928 from MIMIC-III, 1,551 from eICU, and 500 from Guangxi for external validation. AKI was defined using serum creatinine-based Kidney Disease: Improving Global Outcomes criteria. Predictors were collected within the first 24 hours of ICU admission. Three feature-selection methods identified 13 predictors for seven models. Discrimination, calibration, potential clinical utility, and interpretability were evaluated. Logistic regression achieved an area under the receiver operating characteristic curve of 0.817 (95% confidence interval, 0.790–0.844) internally, and 0.807 (0.784–0.830), 0.793 (0.772–0.813), and 0.772 (0.728–0.815) in eICU, MIMIC-III, and Guangxi, respectively. Shapley additive explanations identified blood urea nitrogen, phosphate, chronic kidney disease, lactate, and bicarbonate as the principal contributors to model predictions. The model demonstrated consistent discrimination across cohorts and may support early AKI risk assessment in older ICU patients with aspiration pneumonia, although prospective validation is required before clinical implementation.

1 Introduction

Aspiration pneumonia (AP) is an important form of pulmonary infection in older adults and is commonly associated with dysphagia, neurological disorders, frailty, and multiple chronic comorbidities [1]. Previous guidelines have estimated that approximately 5%–15% of hospitalized pneumonia cases are related to aspiration, with an even higher proportion reported in high-risk populations such as nursing home residents [2]. Among hospitalized older adults, aspiration pneumonia was associated with a 30-day mortality of 31%, compared with 15% in those with non-aspiration pneumonia [3]. Acute kidney injury (AKI) represents a major form of organ dysfunction in critically ill patients and has been linked to both higher short-term mortality and subsequent risks of chronic kidney disease and adverse cardiovascular outcomes [4, 5]. Notably, a study of older patients with community-acquired pneumonia reported an AKI incidence of 34.4%, with an in-hospital mortality rate of 51.6% among patients who developed AKI compared with 14.4% among those without AKI [6]. These findings suggest that renal dysfunction may further contribute to poor outcomes in older patients with pneumonia.

Early recognition of older patients with aspiration pneumonia who are particularly susceptible to AKI remains difficult. This population frequently presents with frailty, multiple comorbidities, and atypical clinical manifestations, while aspiration events themselves may not always be clearly recognized [7]. Moreover, AKI is conventionally identified on the basis of changes in serum creatinine and urine output. However, increases in serum creatinine may lag behind the onset of kidney injury, particularly in older patients with reduced muscle mass, in whom serum creatinine may inadequately reflect early deterioration in renal function. Urine output is also susceptible to the effects of volume status, fluid resuscitation, and diuretic therapy [8, 9]. Therefore, relying solely on established abnormalities in renal function may delay risk recognition, highlighting the potential value of individualized risk assessment based on routinely available clinical information obtained during the early ICU period.

Machine learning approaches can integrate demographic characteristics, comorbidities, vital signs, laboratory measurements, and treatment-related information while capturing complex relationships among variables that may not be fully characterized by conventional statistical methods. These approaches therefore provide a potential strategy for AKI risk stratification. Previous systematic reviews have shown that machine learning models generally achieve favorable discrimination for AKI prediction in ICU populations [10]. The number of AKI prediction models developed across different clinical populations has also increased substantially. One systematic review identified 150 AKI prediction models but reported considerable heterogeneity in study design, AKI definitions, and model evaluation, with most studies considered to be at high risk of bias [11]. More recently, machine learning models have been developed to predict AKI in patients with community-acquired pneumonia and have demonstrated promising performance in both internal and external validation [12]. However, these populations differ substantially from older patients with aspiration pneumonia requiring ICU care. Moreover, external validation remains uncommon in ICU-based artificial intelligence prediction studies. A systematic review found that only 14.7% of such studies performed external validation, with model performance generally declining when evaluated in external datasets [13]. Multicohort validation in this specific high-risk population is therefore warranted.

Accordingly, we constructed and compared several AKI risk models using the MIMIC-IV cohort and subsequently examined their performance in three external validation cohorts, including MIMIC-III, eICU, and a Chinese ICU cohort. The selected model was further evaluated using Shapley additive explanations (SHAP) to characterize predictor contributions. This study aimed to establish an interpretable and transportable risk assessment tool to facilitate AKI risk stratification among older patients with aspiration pneumonia requiring ICU care.

2 Methods

2.1 Data Sources and Study Population

This retrospective multicohort study drew upon three large critical care databases together with an additional ICU cohort from a tertiary hospital in China. The included data sources were MIMIC-IV v3.1, MIMIC-III v1.4 (CareVue subset), the eICU Collaborative Research Database v2.0, and the ICU cohort of the People's Hospital of Guangxi Zhuang Autonomous Region. MIMIC-IV contains deidentified clinical records of hospitalized and critically ill patients treated at Beth Israel Deaconess Medical Center between 2008 and 2022 [14]. MIMIC-III represents an earlier critical care dataset from the same institution, and the CareVue-based records were selected for the present analysis [15]. The eICU Collaborative Research Database comprises multicenter ICU records collected across hospitals in the United States between 2014 and 2015 [16]. For the Guangxi cohort, clinical data were retrieved from the electronic medical record system of the People's Hospital of Guangxi Zhuang Autonomous Region for patients admitted to the ICU between May 2020 and May 2026.

MIMIC-IV served as the derivation and internal evaluation cohort, whereas MIMIC-III, eICU, and the Guangxi cohort were independently used for external validation. The outcome of interest was AKI occurring during the ICU hospitalization, with follow-up extending from ICU admission to ICU discharge.

MIMIC-IV, MIMIC-III, and eICU are publicly accessible databases containing deidentified patient-level records, with access granted to qualified investigators after completion of the required data-access procedures and research ethics training. Qimeng Xu completed the CITI Program training in human research involving data or specimens only (Record ID: 76450437). The collection of patient information and creation of the MIMIC-IV research resource were reviewed by the Institutional Review Board at Beth Israel Deaconess Medical Center (Boston, MA, USA), which granted a waiver of informed consent and approved the data-sharing initiative. The MIMIC-III project was approved by the Institutional Review Boards of Beth Israel Deaconess Medical Center (Boston, MA, USA) and the Massachusetts Institute of Technology (Cambridge, MA, USA), and the requirement for individual patient consent was waived because the project did not affect clinical care and all protected health information was deidentified. The eICU Collaborative Research Database contains deidentified patient data that meet the Safe Harbor provisions of the Health Insurance Portability and Accountability Act (HIPAA), and only these deidentified data were analyzed in the present study. For the Guangxi cohort, the Science and Technology Ethics Committee—Branch of the Ethics Committee, People's Hospital of Guangxi Zhuang Autonomous Region, granted exemption from ethics review based on the use of anonymized information and data (Approval No. KY-IIT-2026-041; August 3, 2026). The requirement to obtain informed consent from participants in the Guangxi cohort or their legal guardians was waived by the same Committee. This retrospective analysis used anonymized clinical data and involved no direct contact with participants. All methods were carried out in accordance with the Declaration of Helsinki and relevant guidelines and regulations, and all patient information analyzed in the present study was deidentified.

2.2 Patient Selection and Definition of AKI

The study population comprised adults aged 65 years or older with aspiration pneumonia who required ICU admission. Cases of aspiration pneumonia were identified from diagnostic coding information within each database using ICD-9-CM code 507.0 and ICD-10-CM code J69.0. Eligibility required a documented diagnosis of aspiration pneumonia, an age of at least 65 years, and an ICU stay of ≥ 24 h. When multiple hospitalizations or ICU encounters were available for the same individual, the first ICU admission satisfying all predefined eligibility criteria was retained, and later encounters were excluded to avoid duplicate enrollment. Individuals aged < 65 years or with an ICU stay of < 24 h were excluded.

AKI was defined using the serum creatinine criteria specified in the 2012 Kidney Disease: Improving Global Outcomes (KDIGO) guideline [17]. The first serum creatinine measurement obtained after ICU admission was used as the reference serum creatinine. AKI was diagnosed when serum creatinine increased by ≥ 0.3 mg/dL (≥ 26.5 μmol/L) within 48 h or increased to ≥ 1.5 times the reference value within 7 days. AKI onset was defined as the first time point at which either criterion was met after ICU admission. Because the availability and completeness of urine output measurements varied substantially across the databases, urine output was not incorporated into the AKI definition to maintain a consistent outcome definition across cohorts.

2.3 Candidate Predictors and Data Preprocessing

Information on clinical variables available during the first 24 h after ICU admission was extracted from the MIMIC-IV, MIMIC-III, and eICU databases. For the Guangxi validation cohort, corresponding information was retrieved from the institutional electronic medical record system according to identical variable definitions and collection windows. Candidate predictors covered demographic characteristics, comorbidities,

physiological measurements, treatment exposures, and laboratory findings. Demographic variables included sex, age, and body weight. The comorbidity variables were cerebrovascular accident (CVA), chronic kidney disease (CKD), type 2 diabetes mellitus (T2DM), chronic obstructive pulmonary disease (COPD), and heart failure (HF). Physiological variables included heart rate, respiratory rate, and body temperature. Treatment-related variables included exposure to gentamicin, vancomycin, and mechanical ventilation. The laboratory panel comprised chloride, potassium, sodium, blood urea nitrogen, hematocrit, anion gap, glucose, bicarbonate, hemoglobin, white blood cell count, platelet count, magnesium, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), red cell distribution width (RDW), red blood cell count, phosphate, total calcium, international normalized ratio (INR), prothrombin time (PT), partial thromboplastin time (PTT), base excess, calculated total carbon dioxide, lactate, partial pressure of carbon dioxide (PCO_2), pH, partial pressure of oxygen (PO_2), and aspartate aminotransferase (AST).

For continuous variables with repeated measurements during the initial 24 h of ICU admission, the earliest recorded measurement within this interval was selected for analysis. Binary variables were encoded according to the presence or absence of the corresponding diagnosis, medication exposure, or therapeutic intervention. The proportion of missing observations was determined for each candidate variable, and variables exceeding a 20% missingness threshold were excluded from subsequent analyses. For the remaining incomplete variables, missing values were handled through multiple imputation using the mice package in R. Following this preprocessing procedure, 43 candidate predictors were retained for subsequent feature selection and model development.

2.4 Feature Selection

To limit unnecessary dimensionality while enhancing model stability and interpretability, feature selection was conducted solely within the MIMIC-IV training set using the available candidate data. Neither the internal test set nor the MIMIC-III, eICU, or Guangxi external validation cohorts contributed to the feature-selection process. Predictor screening incorporated three complementary techniques: least absolute shrinkage and selection operator (LASSO), the Boruta algorithm, and recursive feature elimination (RFE).

Within the LASSO framework, an L1 penalty was used to shrink model coefficients, and the regularization parameter was determined by 10-fold cross-validation. The Boruta procedure was based on a random forest and assessed the importance of each observed predictor relative to randomly generated shadow variables. RFE iteratively removed lower-value predictors and used cross-validation to determine the most appropriate subset. Variables retained simultaneously by all three methods were considered the final predictors for subsequent model development.

The ability of each candidate variable to discriminate between AKI and non-AKI was also assessed individually by calculating its area under the receiver operating characteristic curve (AUC) in the training set. Spearman rank correlation was then used to examine associations between candidate predictors and identify potentially redundant information. A correlation matrix was subsequently generated to examine possible collinearity patterns among predictors.

2.5 Model Development, Validation, and Performance Evaluation

The MIMIC-IV cohort was randomly divided into a training set and an internal validation set at a 7:3 ratio. The training set was used for data preprocessing, feature selection, model fitting, and hyperparameter optimization. Based on the final selected predictors, seven prediction models were developed, including logistic regression, support vector machine (SVM), neural network, k-nearest neighbors (KNN), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), and Light Gradient Boosting Machine (LightGBM). Hyperparameters were optimized using the training set, and the final settings are provided in Supplementary Table S1. The seven candidate models were first evaluated in the MIMIC-IV internal validation set. Model selection was based on their performance in the internal validation set, with consideration of discrimination, calibration, and clinical utility. The model demonstrating the most favorable overall performance in the MIMIC-IV cohort was selected as the final model prior to external validation.

The MIMIC-III, eICU, and Guangxi cohorts were subsequently used as independent external validation datasets to assess the transportability of the final model across different time periods, healthcare systems, and patient populations. Model discrimination was evaluated using receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC). Accuracy, sensitivity, specificity, precision, F1-score, negative predictive value (NPV), and Cohen's kappa coefficient were additionally calculated to characterize classification performance. Calibration curves were used to assess agreement between predicted and observed AKI risks, while decision curve analysis (DCA) was performed to evaluate the potential clinical net benefit across a range of probability thresholds.

2.6 Model Interpretability

Interpretation of the final logistic regression model was performed using Shapley additive explanations (SHAP), which enabled quantification of predictor effects on model-derived risk estimates at both the overall cohort level and the individual-patient level. For the global analysis, predictors were ranked by their mean absolute SHAP values, while SHAP summary plots were used to illustrate the distribution of feature values, the magnitude of their effects, and whether their contributions increased or decreased the predicted probability of AKI. For individual-level interpretation, SHAP explanations were additionally used to characterize the direction and magnitude of predictor-specific effects for individual patients.

SHAP dependence plots were also constructed for predictors making relatively large contributions to model output to further examine how changes in their values were associated with changes in prediction. These analyses further illustrated possible nonlinear relationships in risk estimation and the contribution patterns of the principal predictors related to AKI.

2.7 Statistical Analysis

Continuous variables were described using medians and interquartile ranges [M (Q1, Q3)], according to their underlying distributions, whereas categorical data were expressed as counts and percentages [n (%)]. For comparisons between patients with and without AKI, the Mann–Whitney U test was used for continuous variables, while categorical variables were analyzed with Pearson's χ^2 test or Fisher's exact test, as

appropriate according to expected frequencies and distribution characteristics. All statistical analyses were two-sided, and $P < 0.05$ was considered statistically significant.

R was used as the primary platform for data preprocessing, management of missing observations, statistical analysis, feature selection, model development, and evaluation of predictive performance. Model discrimination was evaluated using ROC curves and AUC values, with additional assessment using accuracy, sensitivity, specificity, precision, F1-score, NPV, and Cohen's kappa coefficient. Calibration was assessed graphically with calibration curves, DCA was used to examine potential clinical utility, and SHAP was employed to investigate model interpretability.

3 Results

3.1 Patient Selection and Baseline Characteristics

Following the predefined eligibility criteria, the final study population consisted of 3,125 patients from MIMIC-IV, 1,928 from MIMIC-III, 1,551 from eICU, and 500 from the Guangxi cohort. The number of patients who developed AKI was 1,615, 775, 406, and 154 across the four cohorts, corresponding to AKI incidences of 51.7%, 40.2%, 26.2%, and 30.8%, respectively. The detailed patient screening procedure is illustrated in Fig. 1.

Table 1 summarizes the baseline characteristics of patients in the MIMIC-IV development cohort. Of the 3,125 included patients, 1,615 experienced AKI during ICU hospitalization, whereas 1,510 remained free of AKI. Patients who developed AKI tended to be older, had higher body weight, and included a greater proportion of males compared with those without AKI. The frequencies of CKD, T2DM, and HF were also increased among patients with AKI. Regarding laboratory findings, the AKI group exhibited elevated levels of blood urea nitrogen, phosphate, potassium, anion gap, lactate, INR, PT, and PTT, accompanied by reduced bicarbonate, total calcium, pH, and PO_2 values, with significant differences observed for most parameters. In addition, vancomycin exposure was more common in patients with AKI, whereas no significant differences were observed between groups for heart rate, CVA, COPD, gentamicin exposure, or mechanical ventilation.

Table 1
Baseline characteristics of older ICU patients with aspiration pneumonia according to AKI status in the MIMIC-IV cohort.

Variable	Total (n = 3,125)	Non-AKI (n = 1,510)	AKI (n = 1,615)	P value
Demographic characteristics				
Age (years)	78 (71, 86)	78 (71, 85)	79 (72, 86)	0.027
Sex				0.004
Female	1198 (38)	618 (41)	580 (36)	
Male	1927 (62)	892 (59)	1035 (64)	
Weight (kg)	73 (61, 85.9)	70.2 (59.28, 82.5)	75.7 (62.92, 89.05)	< 0.001
Vital signs				
Heart rate (bpm)	90 (75, 105)	89 (75, 104.75)	90 (76, 105)	0.613
Respiratory rate (breaths/min)	20 (17, 25)	20 (16, 24)	20 (17, 25)	0.028
Temperature (°C)	36.78 (36.44, 37.17)	36.83 (36.5, 37.17)	36.72 (36.44, 37.11)	< 0.001
Comorbidities				
CVA				0.789
No	2705 (87)	1304 (86)	1401 (87)	
Yes	420 (13)	206 (14)	214 (13)	
CKD				< 0.001
No	2327 (74)	1318 (87)	1009 (62)	
Yes	798 (26)	192 (13)	606 (38)	
T2DM				< 0.001
No	2123 (68)	1118 (74)	1005 (62)	
Yes	1002 (32)	392 (26)	610 (38)	
COPD				1.000
No	2417 (77)	1168 (77)	1249 (77)	

Continuous variables are presented as median (Q1, Q3), and categorical variables as n (%). Abbreviations: AKI, acute kidney injury; ICU, intensive care unit; CVA, cerebrovascular accident; CKD, chronic kidney disease; T2DM, type 2 diabetes mellitus; COPD, chronic obstructive pulmonary disease; HF, heart failure; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red cell distribution width; INR, international normalized ratio; PT, prothrombin time; PTT, partial thromboplastin time; CO_2, carbon dioxide; PCO_2, partial pressure of carbon dioxide; PO_2, partial pressure of oxygen; AST, aspartate aminotransferase.

Variable	Total (n = 3,125)	Non-AKI (n = 1,510)	AKI (n = 1,615)	P value
Yes	708 (23)	342 (23)	366 (23)	
HF				< 0.001
No	1839 (59)	1009 (67)	830 (51)	
Yes	1286 (41)	501 (33)	785 (49)	
Laboratory indicators				
Chloride (mmol/L)	104 (100, 108)	104 (100, 108)	105 (100, 109)	< 0.001
Potassium (mmol/L)	4.1 (3.7, 4.6)	4 (3.6, 4.4)	4.2 (3.8, 4.7)	< 0.001
Sodium (mmol/L)	139 (136, 143)	139 (136, 142)	140 (136, 143)	< 0.001
Blood urea nitrogen (mg/dL)	26 (17, 40)	20 (14, 28)	34 (23, 51)	< 0.001
Hematocrit (%)	32.4 (27.7, 36.8)	32.9 (28.33, 37.1)	32 (27.1, 36.25)	< 0.001
Anion gap (mmol/L)	14 (12, 17)	13 (11, 16)	15 (12, 18)	< 0.001
Glucose (mmol/L)	7.33 (5.89, 9.61)	7.11 (5.78, 9)	7.67 (6.08, 10.22)	< 0.001
Bicarbonate (mmol/L)	23 (20, 26)	24 (21, 27)	22 (19, 25)	< 0.001
Hemoglobin (g/dL)	10.4 (8.8, 11.9)	10.6 (9.1, 12.1)	10.2 (8.6, 11.8)	< 0.001
White blood cell count ($\times 10^3$/μL)	11.7 (8.2, 16.2)	11.4 (8.3, 15.4)	12 (8.1, 16.9)	0.013
Platelet count ($\times 10^3$/μL)	205 (149, 273)	215 (156, 279.75)	198 (142, 266)	< 0.001
Magnesium (mg/dL)	1.9 (1.7, 2.2)	1.9 (1.7, 2.1)	2 (1.7, 2.3)	< 0.001
MCH (pg)	30 (28.4, 31.4)	30.05 (28.5, 31.5)	29.9 (28.35, 31.4)	0.077
MCHC (g/dL)	32.2 (31.1, 33.2)	32.3 (31.3, 33.4)	32 (30.95, 33.1)	< 0.001
MCV (fL)	93 (88, 97)	92 (88, 97)	93 (88, 97)	0.066
RDW (%)	15 (13.8, 16.7)	14.7 (13.7, 16.2)	15.3 (14, 17)	< 0.001
Red blood cell count ($\times 10^6$/μL)	3.5 (3, 4.02)	3.58 (3.06, 4.06)	3.44 (2.94, 3.98)	< 0.001
Phosphate (mg/dL)	3.5 (2.8, 4.2)	3.2 (2.7, 3.8)	3.8 (3.1, 4.7)	< 0.001

Continuous variables are presented as median (Q1, Q3), and categorical variables as n (%). Abbreviations: AKI, acute kidney injury; ICU, intensive care unit; CVA, cerebrovascular accident; CKD, chronic kidney disease; T2DM, type 2 diabetes mellitus; COPD, chronic obstructive pulmonary disease; HF, heart failure; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red cell distribution width; INR, international normalized ratio; PT, prothrombin time; PTT, partial thromboplastin time; CO_2, carbon dioxide; PCO_2, partial pressure of carbon dioxide; PO_2, partial pressure of oxygen; AST, aspartate aminotransferase.

Variable	Total (n = 3,125)	Non-AKI (n = 1,510)	AKI (n = 1,615)	P value
Total calcium (mg/dL)	8.4 (7.9, 8.9)	8.4 (8, 8.9)	8.3 (7.7, 8.8)	< 0.001
INR	1.3 (1.1, 1.5)	1.2 (1.1, 1.4)	1.3 (1.2, 1.7)	< 0.001
PT (s)	14.2 (12.6, 16.9)	13.7 (12.4, 15.97)	14.6 (12.8, 18.1)	< 0.001
PTT (s)	31 (27.4, 37.2)	30.25 (27.1, 35.5)	31.7 (27.9, 38.8)	< 0.001
Base excess (mmol/L)	-1 (-5, 1)	0 (-3, 2)	-2 (-6, 1)	< 0.001
Calculated total CO_2 (mmol/L)	25 (21, 28)	25 (23, 29)	24 (20, 28)	< 0.001
Lactate (mmol/L)	1.7 (1.2, 2.6)	1.6 (1.1, 2.3)	1.8 (1.25, 2.9)	< 0.001
PCO_2 (mmHg)	43 (36, 54)	44 (37, 56)	43 (36, 53)	0.002
pH	7.36 (7.27, 7.42)	7.37 (7.28, 7.43)	7.34 (7.26, 7.41)	< 0.001
PO_2 (mmHg)	78 (46, 127)	82 (49, 140)	74 (44, 118)	< 0.001
AST (U/L)	30 (21, 55)	28 (19, 43)	34 (22, 69.5)	< 0.001
Treatments				
Gentamicin				1.000
No	3084 (99)	1490 (99)	1594 (99)	
Yes	41 (1)	20 (1)	21 (1)	
Vancomycin				< 0.001
No	586 (19)	324 (21)	262 (16)	
Yes	2539 (81)	1186 (79)	1353 (84)	
Mechanical ventilation				0.493
No	312 (10)	157 (10)	155 (10)	
Yes	2813 (90)	1353 (90)	1460 (90)	

Continuous variables are presented as median (Q1, Q3), and categorical variables as n (%). Abbreviations: AKI, acute kidney injury; ICU, intensive care unit; CVA, cerebrovascular accident; CKD, chronic kidney disease; T2DM, type 2 diabetes mellitus; COPD, chronic obstructive pulmonary disease; HF, heart failure; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red cell distribution width; INR, international normalized ratio; PT, prothrombin time; PTT, partial thromboplastin time; CO_2, carbon dioxide; PCO_2, partial pressure of carbon dioxide; PO_2, partial pressure of oxygen; AST, aspartate aminotransferase.

3.2 Feature Selection and Identification of Final Predictors

Feature selection was performed among the 43 candidate predictors using LASSO, Boruta, and RFE. Boruta identified several clinical variables with stable importance across iterations, LASSO selected predictors through cross-validated coefficient shrinkage, and RFE further refined the feature set by comparing predictive

performance across different subsets. The intersection of the three methods yielded 13 final predictors for model development: blood urea nitrogen, phosphate, bicarbonate, CKD, magnesium, HF, body weight, pH, chloride, PTT, PO_2, total calcium, and lactate (Fig. 2). The RFE feature-selection process is shown in Supplementary Figure S1.

Univariable analysis showed that blood urea nitrogen had the highest discriminative ability for AKI, with an AUC of approximately 0.75, followed by phosphate, bicarbonate, and CKD. The remaining candidate variables showed varying degrees of univariable discrimination. Spearman correlation analysis further characterized the correlation structure among the candidate predictors, with varying degrees of correlation observed among several biochemical and physiological variables (Supplementary Figures S2 and S3). Although the ranking of the final 13 predictors differed across algorithms, blood urea nitrogen ranked first in the AdaBoost, KNN, LightGBM, logistic regression, SVM, and XGBoost models, whereas CKD ranked first and blood urea nitrogen second in the neural network model. Feature-importance rankings for all models are presented in Supplementary Figures S4–S10.

3.3 Model Performance and External Validation

Figure 3 presents the ROC curves, calibration plots, and decision curve analyses for the seven models in the MIMIC-IV internal validation cohort and the three external validation cohorts. In the MIMIC-IV internal validation cohort, logistic regression showed the highest discriminatory ability, with an AUC of 0.817 (95% CI, 0.790–0.844), together with favorable calibration and clinical utility. Based on its overall performance in the MIMIC-IV internal validation cohort, logistic regression was selected as the final model for subsequent external validation and interpretation.

The discriminatory performance of logistic regression remained relatively consistent across the external validation cohorts, with AUCs of 0.807 (95% CI, 0.784–0.830) in eICU, 0.793 (95% CI, 0.772–0.813) in MIMIC-III, and 0.772 (95% CI, 0.728–0.815) in the Guangxi cohort. Across the validation datasets, most models demonstrated acceptable discriminative performance, whereas KNN generally showed relatively lower discrimination. Calibration plots showed varying degrees of agreement between model-predicted risks and observed AKI incidence, while DCA indicated potential clinical net benefit across specific ranges of predicted-risk thresholds.

Table 2 summarizes the classification metrics for all evaluated models. In the MIMIC-IV internal validation cohort, logistic regression achieved an accuracy of 0.755, sensitivity of 0.733, specificity of 0.777, precision of 0.779, F1-score of 0.755, NPV of 0.732, and Cohen's kappa coefficient of 0.510. In the eICU, MIMIC-III, and Guangxi external validation cohorts, the corresponding accuracy values for logistic regression were 0.719, 0.737, and 0.732, respectively, with NPV values of 0.913, 0.799, and 0.835. Classification performance varied across algorithms in terms of sensitivity, specificity, and other metrics. Confusion matrices illustrating the classification results for AKI and non-AKI patients across the evaluated models are presented in Supplementary Figures S11–S17.

Table 2
Classification performance of seven prediction models in the internal validation and external validation cohorts.

Dataset	Model	Accuracy	Sensitivity	Specificity	Precision	F1-score	NPV	Kappa
MIMIC-IV test set	Logistic regression	0.755	0.733	0.777	0.779	0.755	0.732	0.510
	SVM	0.752	0.738	0.768	0.773	0.755	0.733	0.505
	Neural network	0.744	0.733	0.755	0.762	0.747	0.726	0.488
	KNN	0.658	0.554	0.770	0.720	0.626	0.618	0.322
	XGBoost	0.729	0.750	0.706	0.732	0.741	0.726	0.457
	AdaBoost	0.710	0.705	0.715	0.726	0.715	0.694	0.419
	LightGBM	0.713	0.653	0.777	0.758	0.701	0.677	0.428
eICU validation cohort	Logistic regression	0.719	0.815	0.685	0.478	0.603	0.913	0.407
	SVM	0.741	0.759	0.735	0.504	0.606	0.896	0.425
	Neural network	0.742	0.732	0.746	0.505	0.598	0.887	0.417
	KNN	0.593	0.581	0.597	0.338	0.428	0.801	0.144
	XGBoost	0.694	0.766	0.668	0.450	0.567	0.890	0.354
	AdaBoost	0.646	0.734	0.615	0.403	0.521	0.867	0.276
	LightGBM	0.689	0.653	0.702	0.437	0.524	0.851	0.306
MIMIC-III validation cohort	Logistic regression	0.737	0.720	0.748	0.657	0.687	0.799	0.460
	SVM	0.743	0.654	0.802	0.690	0.672	0.775	0.460
	Neural network	0.719	0.742	0.704	0.628	0.680	0.802	0.433
	KNN	0.707	0.568	0.801	0.657	0.609	0.734	0.377
	XGBoost	0.724	0.683	0.752	0.649	0.665	0.779	0.431
	AdaBoost	0.667	0.774	0.595	0.562	0.651	0.797	0.348
	LightGBM	0.700	0.795	0.637	0.595	0.681	0.822	0.409

Abbreviations: NPV, negative predictive value; SVM, support vector machine; KNN, k-nearest neighbors; XGBoost, extreme gradient boosting; AdaBoost, adaptive boosting; LightGBM, Light Gradient Boosting Machine.

Dataset	Model	Accuracy	Sensitivity	Specificity	Precision	F1-score	NPV	Kappa
Guangxi validation cohort	Logistic regression	0.732	0.662	0.763	0.554	0.604	0.835	0.404
	SVM	0.760	0.656	0.806	0.601	0.627	0.840	0.451
	Neural network	0.714	0.669	0.734	0.528	0.590	0.833	0.375
	KNN	0.626	0.532	0.668	0.416	0.467	0.762	0.186
	XGBoost	0.660	0.727	0.630	0.467	0.569	0.838	0.309
	AdaBoost	0.610	0.857	0.500	0.433	0.575	0.887	0.281
	LightGBM	0.688	0.636	0.711	0.495	0.557	0.815	0.322

Abbreviations: NPV, negative predictive value; SVM, support vector machine; KNN, k-nearest neighbors; XGBoost, extreme gradient boosting; AdaBoost, adaptive boosting; LightGBM, Light Gradient Boosting Machine.

3.4 Model Interpretation and Major Predictors

The final logistic regression model was interpreted using SHAP analysis to quantify the contribution of individual predictors. Global SHAP analysis identified blood urea nitrogen, phosphate, CKD, lactate, and bicarbonate as the variables with the greatest influence on AKI risk estimation, followed by chloride, body weight, total calcium, PO_2, HF, pH, PTT, and magnesium (Fig. 4A, B). Higher levels of blood urea nitrogen, phosphate, and lactate, as well as the presence of CKD, were generally associated with positive SHAP contributions, whereas increased bicarbonate levels mainly contributed negatively to model output. Individual-level SHAP analyses further demonstrated that predictors influenced patient-specific risk estimates with different directions and magnitudes. For a representative patient classified as high risk, elevated blood urea nitrogen and phosphate levels together with CKD contributed substantially to increased predicted risk. Conversely, in a representative low-risk patient, lower blood urea nitrogen and phosphate levels and the absence of CKD contributed to a lower model-estimated risk (Fig. 4C, D).

SHAP dependence analyses further characterized the relationships between major predictors and their contributions to model output (Fig. 5). SHAP values generally increased with increasing blood urea nitrogen, phosphate, and lactate levels. The contribution of blood urea nitrogen increased most markedly across the lower-to-moderate range before gradually plateauing, while patients with CKD showed overall higher positive SHAP contributions than those without CKD. In contrast, bicarbonate showed an overall inverse relationship with predicted AKI risk, whereas increasing chloride levels were accompanied by progressively higher SHAP values. SHAP dependence patterns for the remaining predictors, including total calcium, HF, magnesium, pH, PO_2, PTT, and body weight, are shown in Supplementary Figures S18–S24.

4 Discussion

Using four independent critical care cohorts, we developed and externally validated an AKI risk prediction model specifically for older ICU patients with aspiration pneumonia. Thirteen predictors were identified through combined feature selection using LASSO, Boruta, and RFE, and seven different algorithms were subsequently compared. Logistic regression provided consistent discrimination in the MIMIC-IV internal test set and across the eICU, MIMIC-III, and Guangxi external cohorts, while also showing favorable calibration and potential clinical net benefit. According to the SHAP analysis, blood urea nitrogen, phosphate, CKD, lactate, and bicarbonate exerted the greatest influence on model predictions. Notably, despite the inclusion of several more complex algorithms, the comparatively simple logistic regression approach remained robust across the different cohorts, indicating that increasing algorithmic complexity does not necessarily improve predictive accuracy or model transportability.

The predictors emerging from our analysis were generally in agreement with previously reported AKI-related clinical factors. Blood urea nitrogen (BUN) showed the strongest predictive contribution in our model. Previous investigations in ICU populations have likewise linked higher BUN-related measures to an elevated likelihood of AKI [18]. Elevated BUN may reflect not only impaired renal clearance but also reduced renal perfusion and enhanced catabolic activity during critical illness, and may therefore capture both overall disease burden and renal vulnerability. CKD also emerged as a prominent predictor in our model. A systematic review focusing on older populations reported that pre-existing CKD was strongly associated with an increased risk of AKI [19], while White et al. further showed across 12 ICUs that patients with sepsis and underlying CKD were more likely to develop AKI and had poorer renal recovery and clinical outcomes [20]. One possible explanation is that pre-existing CKD reduces renal functional reserve, thereby increasing susceptibility to infection-related and hemodynamic stress, particularly among older individuals [21]. Our findings also linked higher phosphate and lactate concentrations, together with lower bicarbonate levels, to greater predicted AKI risk. Fang et al. similarly observed an association between elevated serum phosphate and sepsis-associated AKI in both MIMIC-IV and eICU [22]. A prospective ICU cohort study reported that lower bicarbonate levels were associated with AKI occurrence and poorer renal recovery [23], while Zhang et al., in an analysis of 11,431 patients with sepsis, found that higher lactate levels were independently associated with both the occurrence and severity of AKI [24]. Collectively, these predictors may represent complementary dimensions of AKI susceptibility, encompassing renal reserve, tissue perfusion, acid-base balance, and metabolic abnormalities [25].

In recent years, several studies have explored machine learning approaches for AKI prediction in patients with pneumonia. Ma et al. evaluated several machine learning algorithms in patients with community-acquired pneumonia and conducted independent external validation, supporting the feasibility of such approaches for pneumonia-related AKI risk stratification; however, their study population mainly consisted of general hospitalized patients rather than older ICU patients with aspiration pneumonia [12]. Ji et al. later constructed eight machine learning models in critically ill patients with community-acquired pneumonia, but the analysis relied predominantly on the MIMIC-IV database and did not include independent validation across different healthcare systems or national populations [26]. Recent systematic reviews have also highlighted substantial heterogeneity in AKI prediction studies with respect to patient populations, outcome definitions, and modeling strategies, while external validation remains uncommon among ICU artificial intelligence models [13, 27]. In comparison, our study specifically addressed older ICU patients with

aspiration pneumonia and incorporated three independent external cohorts from MIMIC-III, eICU, and a Chinese population, thereby providing validation across different periods, institutions, healthcare settings, and national contexts. We also evaluated model performance beyond discrimination by incorporating calibration, DCA, and SHAP-based interpretation. Accordingly, the principal contribution of the present study is less related to the adoption of increasingly sophisticated algorithms than to the clearly defined high-risk population and the breadth of cross-cohort external validation.

Several limitations of this study warrant consideration. First, owing to the retrospective nature of the study, potential selection bias, inaccuracies in recorded data, and residual confounding factors cannot be completely excluded. Second, because aspiration pneumonia was primarily ascertained from diagnostic codes, differences in recognition of aspiration events and documentation practices across databases and institutions could have resulted in some degree of case misclassification. Third, to maintain consistency in AKI assessment across cohorts, only the KDIGO serum creatinine criteria were applied, without incorporation of urine output measurements. Moreover, the first serum creatinine measurement obtained after ICU admission was adopted as the reference value, which may have led to misclassification of certain AKI events. Fourth, candidate predictors were collected during the first 24 h after ICU admission, whereas AKI monitoring began at ICU admission itself. Consequently, in some patients, predictor ascertainment and outcome assessment may have overlapped temporally, potentially allowing early AKI-related information to be incorporated into the model predictors. Therefore, the proposed model should be interpreted primarily as an early risk-stratification tool for use during ICU care rather than as a strictly prospective prediction model in which every predictor is measured before AKI onset. Furthermore, treatment-related variables may reflect both disease severity and clinician decision-making processes, introducing potential confounding by indication. Finally, although three independent cohorts were used for external validation, the Chinese cohort was derived from a single institution and contained a comparatively limited number of patients. Larger multicenter prospective investigations will therefore be necessary to establish the robustness and clinical applicability of the model in routine practice.

5 Conclusion

Across four independent ICU cohorts, we developed and externally validated an interpretable AKI risk model for older patients with aspiration pneumonia. Logistic regression demonstrated stable discrimination, calibration, and potential clinical utility across three external validation cohorts, with blood urea nitrogen, phosphate, CKD, lactate, and bicarbonate emerging as the principal contributors to model prediction. The model may support early AKI risk stratification, although prospective multicenter validation is required before clinical implementation.

Declarations

Data availability

The MIMIC-IV database, MIMIC-III CareVue subset, and eICU Collaborative Research Database that support the findings of this study are available through PhysioNet to credentialed researchers who complete the required training and comply with the applicable data use agreements. The data from the Guangxi cohort are

not publicly available due to patient privacy and institutional restrictions. These data are, however, available from the corresponding authors upon reasonable request and with the approval of the relevant institution, subject to applicable data protection requirements.

Author contributions

H.L., Q.X., and Y.L. contributed to the conception and design of the study. H.L. and Y.L. collected and curated the clinical data. Q.X. extracted data from the public databases, performed the statistical analyses, and developed and validated the prediction models. S.T., Y.H., and J.Q. contributed to data curation, verification, and interpretation. H.L. and Q.X. drafted the manuscript, and Y.L. contributed to manuscript preparation and revision. S.X. provided clinical guidance and critically revised the manuscript. X.L. supervised and coordinated the study and critically revised the manuscript. All authors reviewed and approved the final manuscript and agree to be accountable for all aspects of the work. H.L., Q.X., and Y.L. contributed equally to this work.

Competing interests

The authors declare no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors thank the teams responsible for developing and maintaining the MIMIC-IV, MIMIC-III, and eICU Collaborative Research Database for making these resources available for research.

References

1. Mandell, L. A., Niederman, M. S. & Aspiration Pneumonia *N Engl. J. Med.* **380**, 651–663, doi:10.1056/NEJMra1714562. (2019).
2. Olson, G. & Davis, A. M. Diagnosis and Treatment of Adults With Community-Acquired Pneumonia. *JAMA* **323**, 885. 10.1001/jama.2019.21118 (2020).
3. Putot, A., Putot, S. & Manckoundia, P. Long-Term Survival After Aspiration Pneumonia in Older Inpatients: A Comparative Study. *J. Am. Med. Dir. Assoc.* **24**, 1088–1091. 10.1016/j.jamda.2023.04.014 (2023).
4. Bellomo, R., Kellum, J. A. & Ronco, C. Acute Kidney Injury. *Lancet* **380**, 756–766. 10.1016/S0140-6736(11)61454-2 (2012).
5. Pickkers, P. et al. Acute Kidney Injury in the Critically Ill: An Updated Review on Pathophysiology and Management. *Intensive Care Med.* **47**, 835–850. 10.1007/s00134-021-06454-7 (2021).
6. Lubart, E., Boguslavsky, T., Goltsman, G., Muhtaseb, S. & Matveychuk, A. The Incidence of Acute Renal Failure and High Mortality Rate in Elderly Patients Hospitalized with Community Acquired Pneumonia. *Exp. Gerontol.* **179**, 112242. 10.1016/j.exger.2023.112242 (2023).

7. Yoshimatsu, Y., Melgaard, D., Westergren, A., Skrubbeltrang, C. & Smithard, D. G. The Diagnosis of Aspiration Pneumonia in Older Persons: A Systematic Review. *Eur. Geriatr. Med.* **13**, 1071–1080. 10.1007/s41999-022-00689-3 (2022).
8. Rossiter, A., La, A., Koyner, J. L. & Forni, L. G. New Biomarkers in Acute Kidney Injury. *Crit. Rev. Clin. Lab. Sci.* **61**, 23–44. 10.1080/10408363.2023.2242481 (2024).
9. Ostermann, M. et al. Biomarkers in Acute Kidney Injury. *Ann. Intensiv. Care.* **14**, 145. 10.1186/s13613-024-01360-9 (2024).
10. Du, Y. H., Guan, C. J., Li, L. Y. & Gan, P. Predictive Value of Machine Learning for the Risk of Acute Kidney Injury (AKI) in Hospital Intensive Care Units (ICU) Patients: A Systematic Review and Meta-Analysis. *PeerJ* **11**, e16405. 10.7717/peerj.16405 (2023).
11. Feng, Y. et al. Characterization of Risk Prediction Models for Acute Kidney Injury: A Systematic Review and Meta-Analysis. *JAMA Netw. Open.* **6**, e2313359. 10.1001/jamanetworkopen.2023.13359 (2023).
12. Ma, M. et al. A Machine Learning-Based Prediction Model for Acute Kidney Injury in Patients With Community-Acquired Pneumonia: Multicenter Validation Study. *J. Med. Internet Res.* **26**, e51255. 10.2196/51255 (2024).
13. Rockenschaub, P. et al. External Validation of AI-Based Scoring Systems in the ICU: A Systematic Review and Meta-Analysis. *BMC Med. Inf. Decis. Mak.* **25** (5). 10.1186/s12911-024-02830-7 (2025).
14. Johnson, A. E. W. et al. MIMIC-IV, a Freely Accessible Electronic Health Record Dataset. *Sci. Data*. **10**, 1. 10.1038/s41597-022-01899-x (2023).
15. Johnson, A. E. W. et al. MIMIC-III, a Freely Accessible Critical Care Database. *Sci. Data*. **3** 10.1038/sdata.2016.35 (2016).
16. Pollard, T. J. et al. The eICU Collaborative Research Database, a Freely Available Multi-Center Database for Critical Care Research. *Sci. Data*. **5**, 180178. 10.1038/sdata.2018.178 (2018).
17. Kellum, J. A., Lameire, N. & for the KDIGO AKI Guideline Work Group. Diagnosis, Evaluation, and Management of Acute Kidney Injury: A KDIGO Summary (Part 1). *Crit. Care*. **17** 10.1186/cc11454 (2013).
18. Yang, F. et al. Prognostic Value of Blood Urea Nitrogen to Serum Albumin Ratio for Acute Kidney Injury and In-Hospital Mortality in Intensive Care Unit Patients with Intracerebral Haemorrhage: A Retrospective Cohort Study Using the MIMIC-IV Database. *BMJ Open.* **13**, e069503. 10.1136/bmjopen-2022-069503 (2023).
19. Stille, K., Kribben, A., Herget-Rosenthal, S. & Incidence Severity, Risk Factors and Outcomes of Acute Kidney Injury in Older Adults: Systematic Review and Meta-Analysis. *J. Nephrol.* **35**, 2237–2250. 10.1007/s40620-022-01381-2 (2022).
20. White, K. C. et al. Sepsis-Associated Acute Kidney Injury in Patients with Chronic Kidney Disease: Patient Characteristics, Prevalence, Timing, Trajectory, Treatment and Associated Outcomes. *Nephrol. (Carlton)*. **29**, 838–848. 10.1111/nep.14392 (2024).
21. Yamamoto, T. & Isaka, Y. Pathological Mechanisms of Kidney Disease in Ageing. *Nat. Rev. Nephrol.* **20**, 603–615. 10.1038/s41581-024-00868-4 (2024).
22. Fang, Y., Zhang, Y. & Zhang, X. Serum Phosphate Levels and the Development of Sepsis Associated Acute Kidney Injury: Evidence from Two Independent Databases. *Front. Med.* **11**, 1367064.

10.3389/fmed.2024.1367064 (2024).

23. Ribeiro, H. S., Duarte, M. P., Burdmann, E. A., Ferreira, A. P. & Inda-Filho, A. J. Serum Bicarbonate Levels and Kidney Outcomes in Critically Ill Patients: A Prospective Cohort Study. *Int. Urol. Nephrol.* **56**, 2983–2989. 10.1007/s11255-024-04029-1 (2024).
24. Zhang, S., Luo, M., Lu, Z. & Shi, Q. Lactate and Lactylation in Sepsis-Associated Acute Kidney Injury: Clinical Evidence from the MIMIC-IV Database and Mechanistic Insights. *Front. Med. (Lausanne)*. **12**, 1708145. 10.3389/fmed.2025.1708145 (2025).
25. Zarbock, A. et al. Sepsis-Associated Acute Kidney Injury: Consensus Report of the 28th Acute Disease Quality Initiative Workgroup. *Nat. Rev. Nephrol.* **19**, 401–417. 10.1038/s41581-023-00683-3 (2023).
26. Ji, W. et al. Prediction of Acute Kidney Injury in Critically Ill Patients with Community-Acquired Pneumonia Using Machine Learning. *J. Intensive Care Med.* **40**, 1247–1259. 10.1177/08850666251349792 (2025).
27. Cama-Olivares, A. et al. Systematic Review and Meta-Analysis of Machine Learning Models for Acute Kidney Injury Risk Classification. *J. Am. Soc. Nephrol.* **36**, 1969–1983. 10.1681/ASN.0000000702 (2025).

Figures

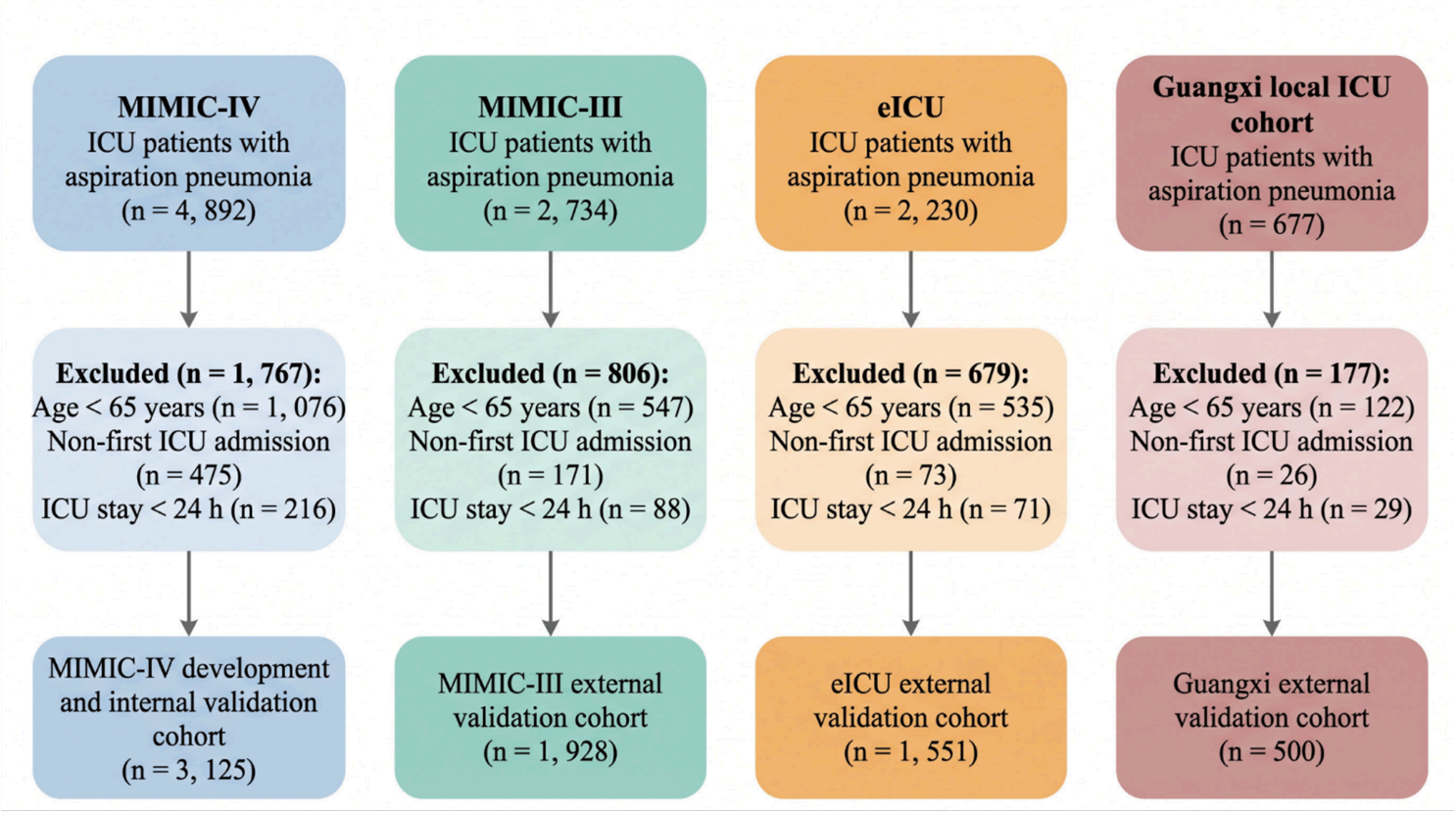


Figure 1

Flowchart of patient selection.Older ICU patients with aspiration pneumonia were screened from the MIMIC-IV, MIMIC-III, eICU, and Guangxi cohorts according to the predefined inclusion and exclusion criteria. The MIMIC-IV cohort was used for model development and internal validation, whereas the MIMIC-III, eICU, and

Guangxi cohorts were used for external validation. Abbreviations:ICU, intensive care unit; MIMIC, Medical Information Mart for Intensive Care; eICU, eICU Collaborative Research Database.

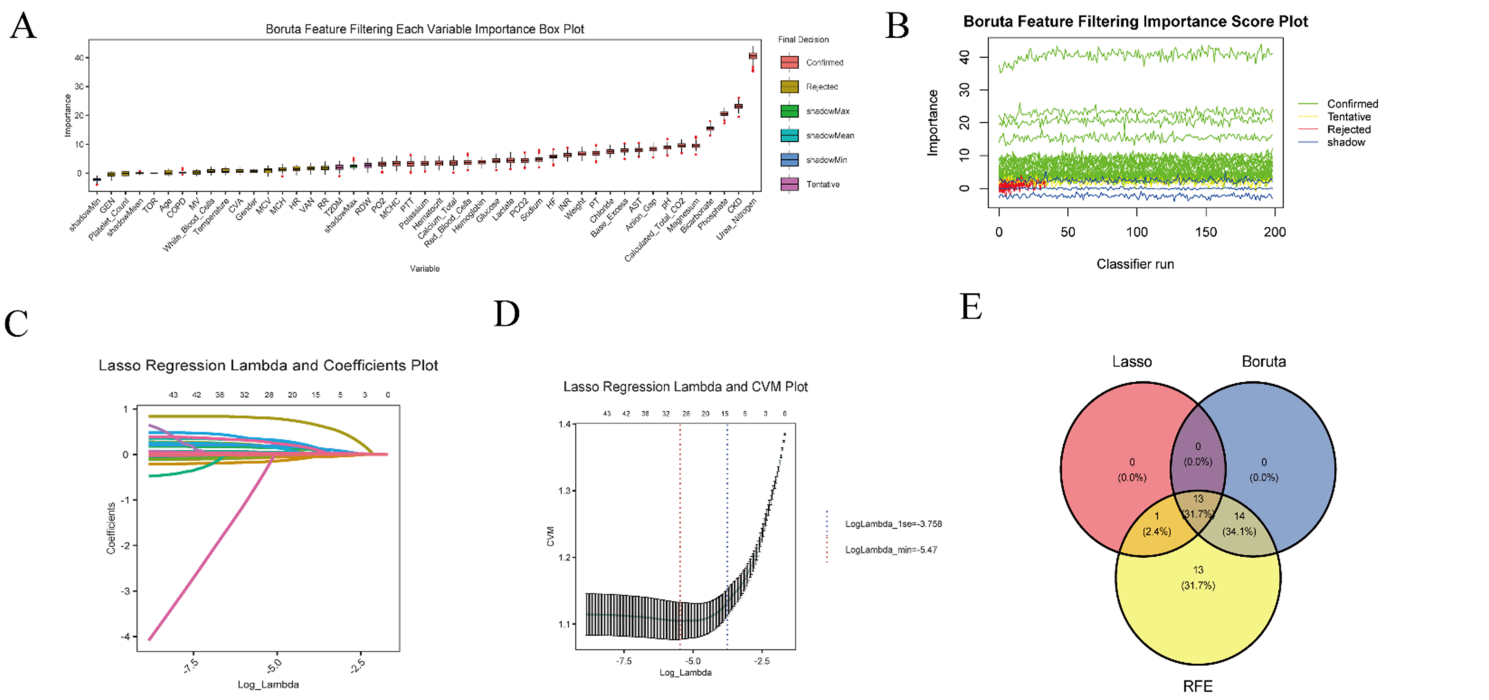


Figure 2

Feature selection using Boruta, LASSO, and recursive feature elimination. (A) Distribution of feature importance assessed by the Boruta algorithm. (B) Evolution of feature importance across Boruta iterations. (C) LASSO coefficient profiles across different values of the regularization parameter. (D) Cross-validation for selection of the optimal LASSO regularization parameter. (E) Venn diagram showing the overlap of predictors selected by LASSO, Boruta, and recursive feature elimination, with 13 predictors retained by all three methods. Abbreviations: LASSO, least absolute shrinkage and selection operator; RFE, recursive feature elimination.

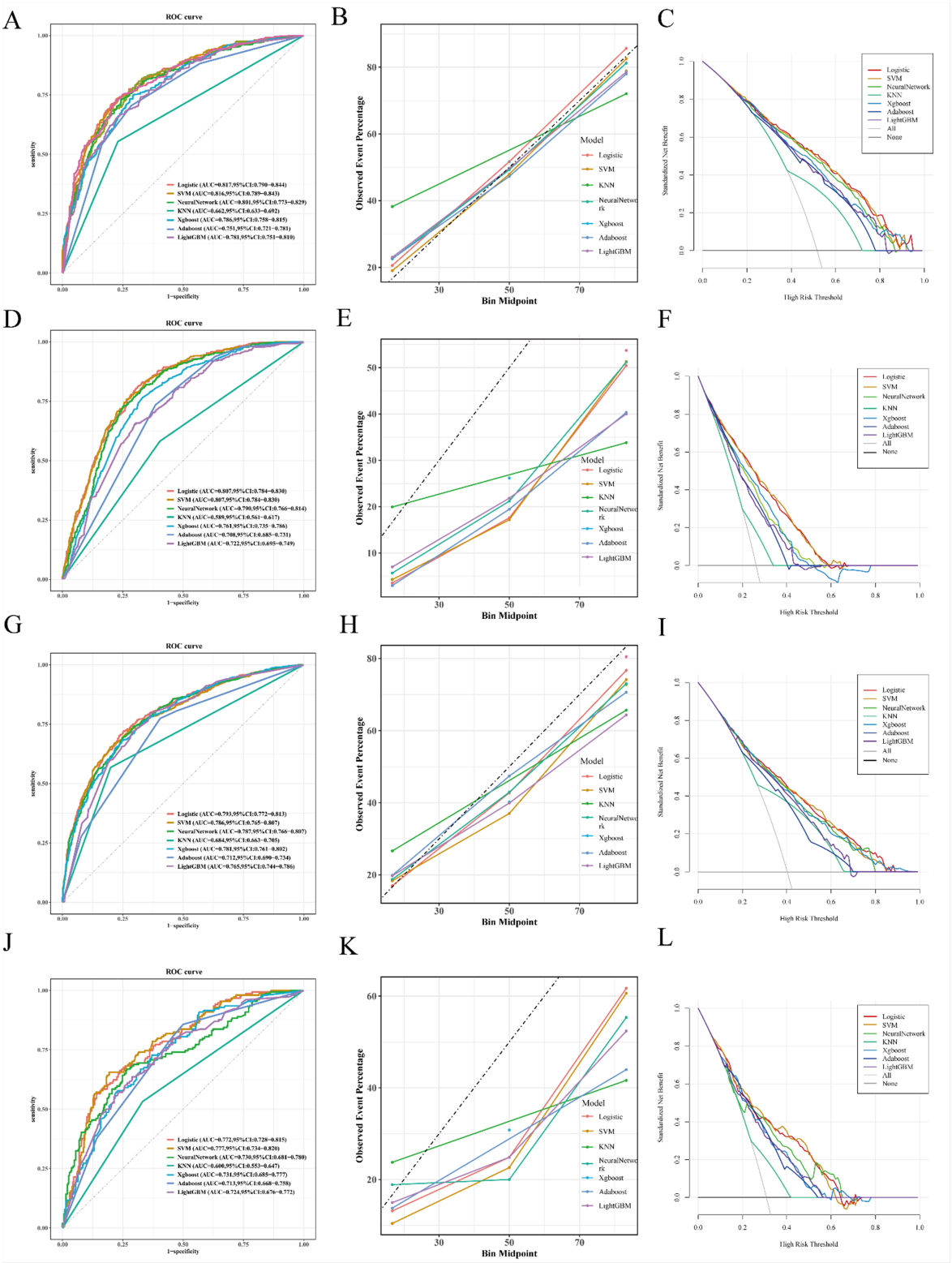


Figure 3

Predictive performance of seven prediction models in the internal validation and external validation cohorts. (A–C) Receiver operating characteristic curve, calibration curve, and decision curve analysis in the MIMIC-IV internal validation cohort. (D–F) Corresponding analyses in the eICU external validation cohort. (G–I) Corresponding analyses in the MIMIC-III external validation cohort. (J–L) Corresponding analyses in the Guangxi external validation cohort. Abbreviations:ROC, receiver operating characteristic; DCA, decision curve

analysis; SVM, support vector machine; KNN, k-nearest neighbors; XGBoost, extreme gradient boosting; AdaBoost, adaptive boosting; LightGBM, Light Gradient Boosting Machine.

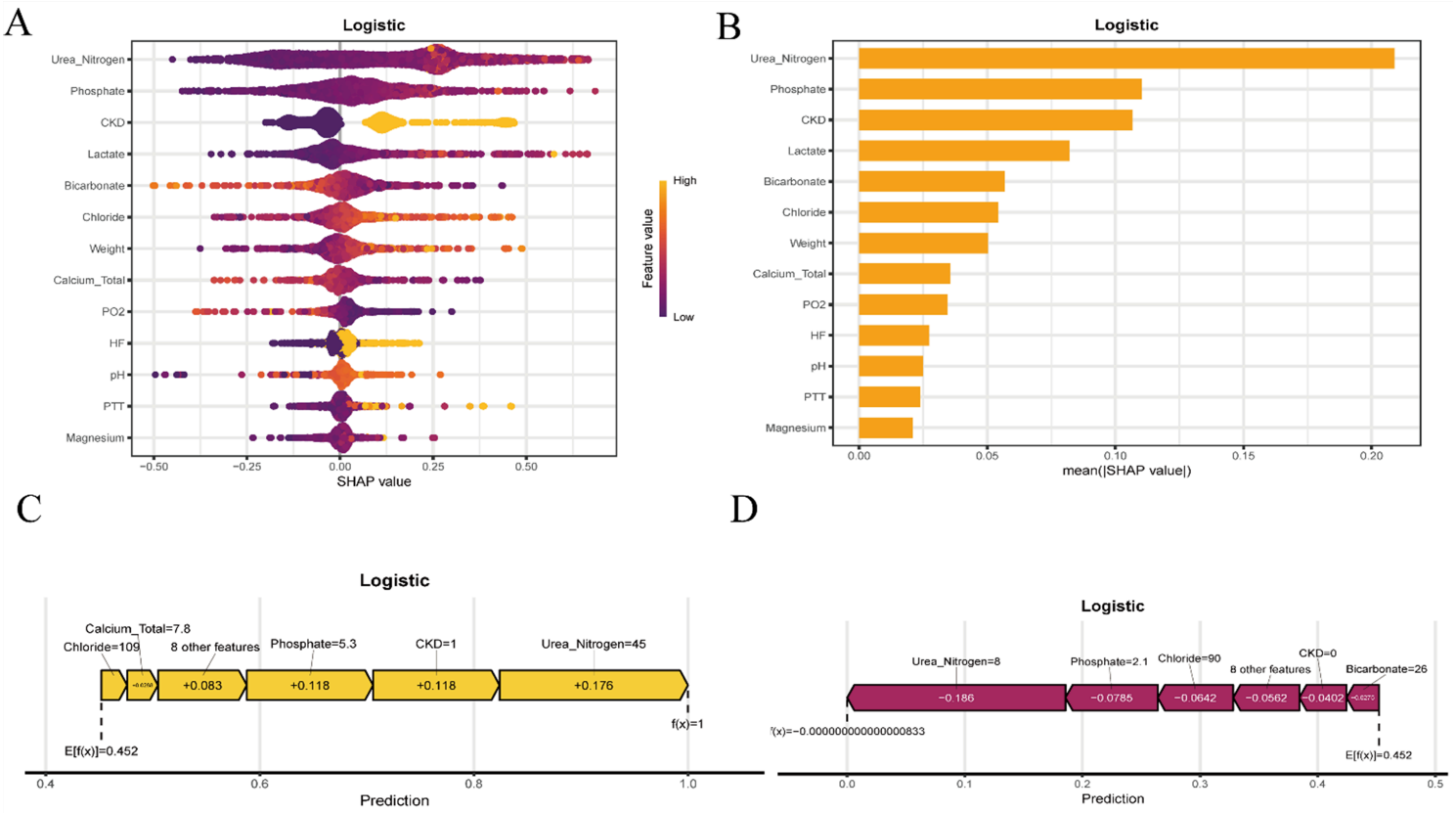


Figure 4

SHAP-based interpretation of the final logistic regression model. (A) SHAP summary plot showing the direction and magnitude of predictor contributions to AKI risk estimates. (B) Ranking of predictors according to mean absolute SHAP values. (C) Individual SHAP explanation for a representative high-risk patient. (D) Individual SHAP explanation for a representative low-risk patient. Abbreviations: SHAP, Shapley additive explanations; AKI, acute kidney injury; CKD, chronic kidney disease; HF, heart failure; PTT, partial thromboplastin time; PO_2, partial pressure of oxygen.

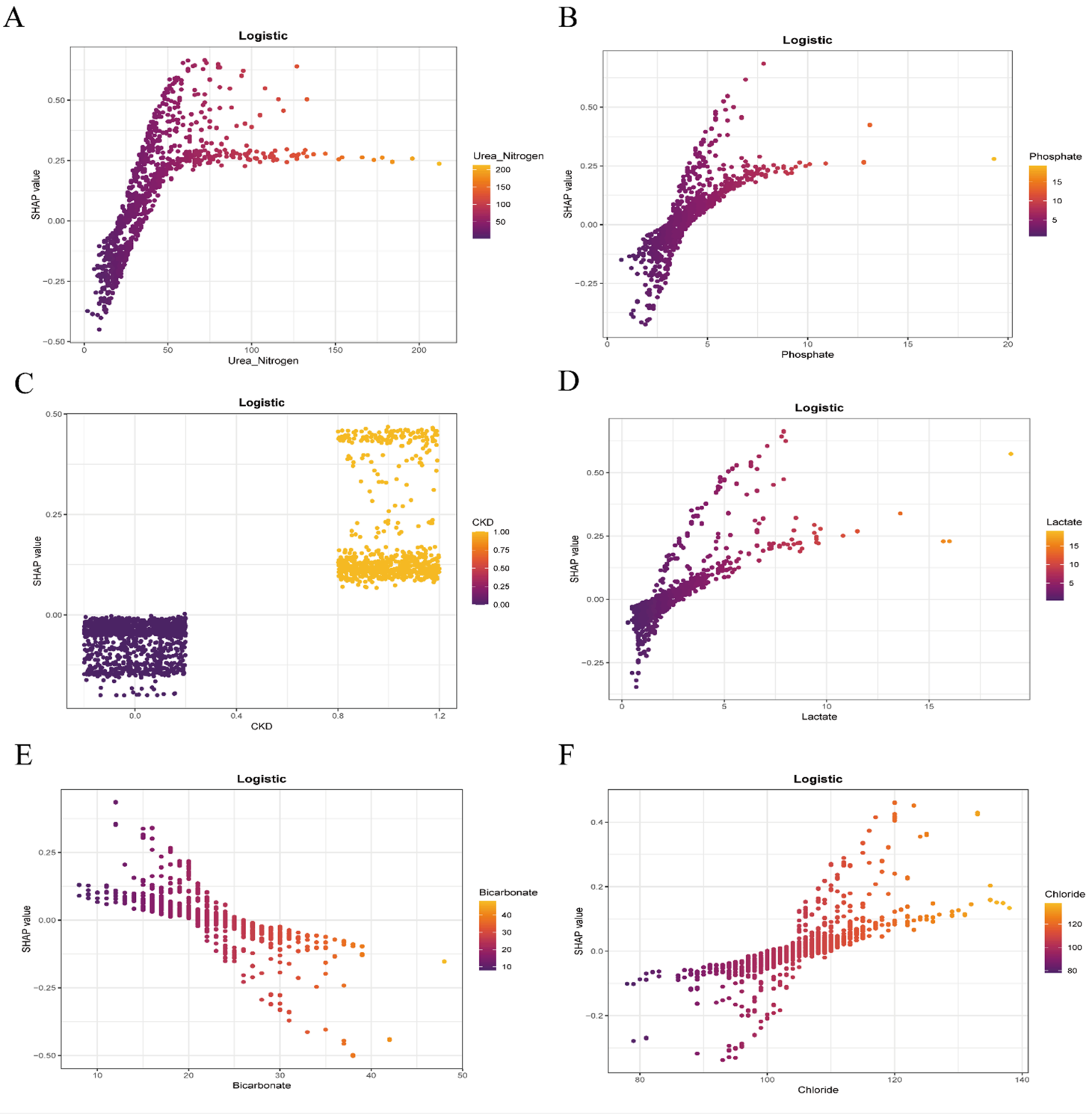


Figure 5

SHAP dependence plots of the major predictors in the final logistic regression model.(A) Blood urea nitrogen. (B) Phosphate. (C) Chronic kidney disease. (D) Lactate. (E) Bicarbonate. (F) Chloride. Abbreviations:SHAP, Shapley additive explanations; CKD, chronic kidney disease.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.docx