Article

Discovery Medicine 2023; 35(177): 623–631
https://doi.org/10.24976/Discov.Med.202335177.61

Development and Comparative Analysis of an Early Prediction Model for Acute Kidney Injury within 72-Hours Post-ICU Admission Using Evidence from the MIMIC-III Database

Yan Luo[1,2], Wenling Ye[3], Yawei Sun[2], Heling Bao[1], Hui Liu[1,*]

[1]Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, 100020 Beijing, China
[2]School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, 100876 Beijing, China
[3]Department of Nephrology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, 100730 Beijing, China
*Correspondence: liuhui@pumc.edu.cn (Hui Liu)
Published: 1 August 2023

Background: Prompt recognition of patients predisposed to acute kidney injury (AKI) within 72 hours of intensive care unit (ICU) admission holds significant clinical importance as it can considerably lower mortality rates. However, existing AKI prediction models often require complex data collection yet yield only moderate performance. This study aims to develop a straightforward and efficient AKI prediction model, providing ICU physicians with a powerful tool to expedite the detection of AKI patients.
Methods: This study proposed a novel generative adversarial imputation networks-least absolute shrinkage and selection operator-extreme gradient boosting (Gain-Lasso-XGBoost) framework and developed an AKI prediction model on the basis of the medical information mart for intensive care (MIMIC-III) database. All the steps, including data preprocessing, feature selection, development, and optimization of prediction models, are organically integrated into the framework which has strong scalability. To compare the performance of our model with current models, we conducted a systematic review to collect all studies on the basis of the MIMIC-III database with similar objectives.
Results: From 15 demographic and clinical variables, 8 features and 5 features were identified as the optimal group of features and processed into the model development. The model optimization further improved the performance of our proposed framework, and the area under curve (AUC) results with 8 and 5 feature vectors achieved 0.849 and 0.830, respectively. Compared with other studies, our method extracted only 8 or 5 feature vectors and obtained superior performance, with an average AUC 1.9% higher than the state-of-the-art approaches in the same type.
Conclusions: Our study suggested that the onset of AKI be effectively and quickly predicted using simplified features, and not just for more specific patient groups. It may help clinicians accurately identify patients at risk of AKI after ICU admission and provide timely monitoring and treatment.

Keywords: acute kidney injury; intensive care unit; model interpretation; clinical decision support; prediction model

Introduction

The intensive care unit (ICU) stands out as a unique department within the hospital, requiring specialized monitoring equipment and the dedicated attention of numerous medical professionals. Moreover, the cost associated with ICU care is typically 3.5 times higher than that of standard patient care [1]. Acute kidney injury (AKI) is a prevalent condition, representing 22–57% of cases among patients in the ICU [2]. AKI can result in extended hospital stays, escalating healthcare costs, and precipitating severe patient medical complications, even leading to life-threatening outcomes [3]. Due to unrecognized kidney injury and delayed diagnosis, AKI has a consistently high mortality rate, accounting for 60% of ICU patients [4]. In this situation, early prediction and active prevention for AKI patients in overwhelmed ICU systems are vital so that medical resources can be proactively allocated during ICU stays.

The development of AKI risk prediction models has flourished in recent years, and researchers found that AKI predictors allowed for the selection of high-risk patients or reduced false positives, and achieved certain results in the early predictions of AKI, which could provide its prediction earlier than physicians [5]. Unfortunately, there are two main limitations in these prior AKI prediction methods. First, most of the studies have used predictors that are specific to certain types of patients (e.g., burn) [6] or procedures (e.g., surgery) [7] and do not generalize to other

Copyright: © 2023 The Author(s). Published by Discovery Medicine. This is an open access article under the CC BY 4.0 license.
Note: *Discovery Medicine* stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

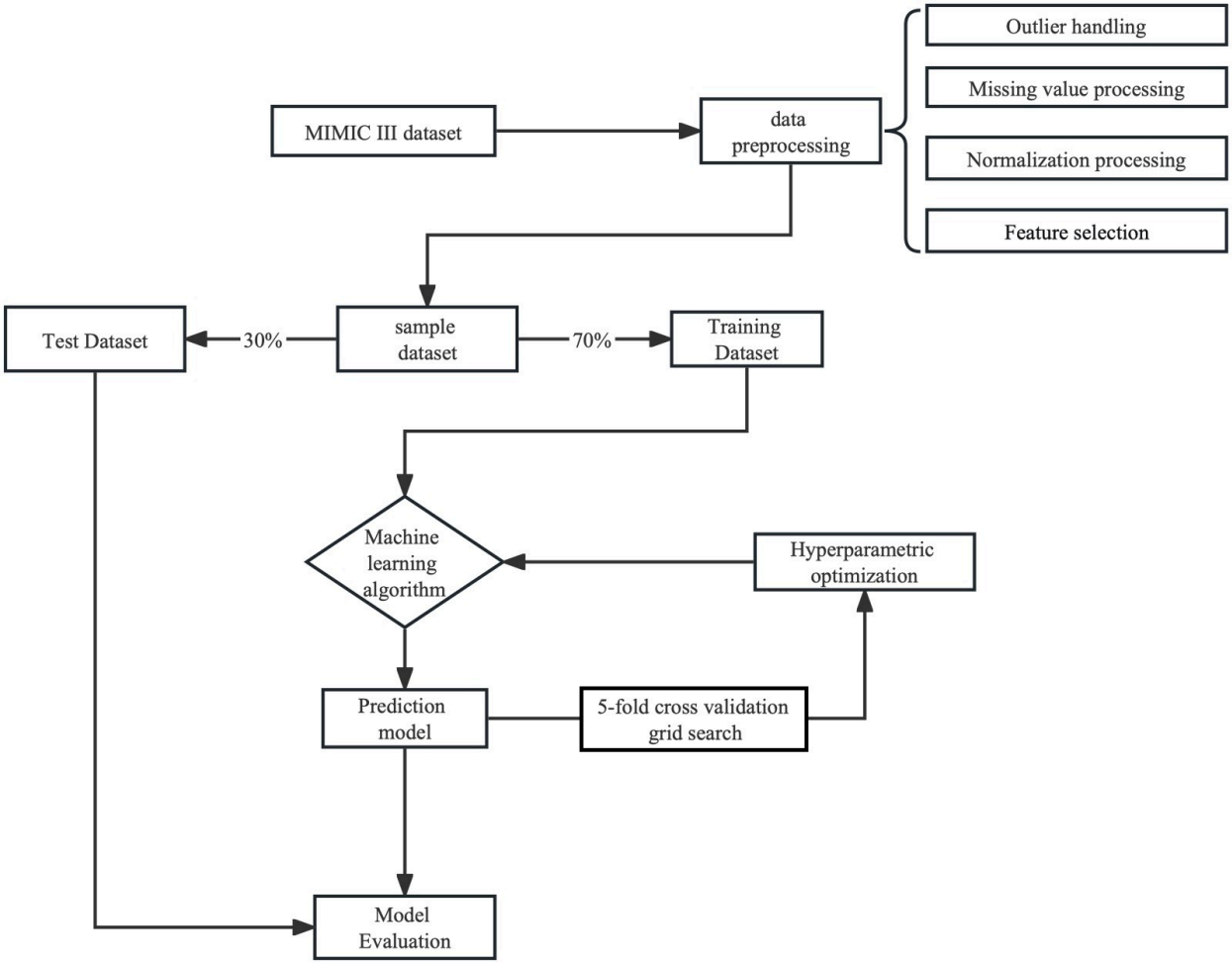

Fig. 1. Overall design of the proposed framework.

prediction problems. Furthermore, most existing prediction models necessitate numerous features for modeling (such as comorbidities, medications, etc.), resulting in complexity and reduced practical applicability. To address this problem, we introduce a new framework to achieve early AKI predictions by using a few variables. To the best of our knowledge, in prior research, scholars do not report the significance of reducing features in predictive models for ICU settings.

Materials and Methods

Data Sources

The retrospective cohort used in this study was selected from a publicly available, large-scale database, the medical information mart for intensive care (MIMIC-III). The Massachusetts Institute of Technology Lab for Computation Physiology developed this critical care dataset to support research in intelligent patient monitoring. It integrated anonymized, comprehensive clinical data with 46,520 ICU cases for all consecutive adult patients admitted to the Beth Israel Deaconess Medical Center between 2001 and 2012 [8].

This study aimed to predict the occurrence of stage 1 AKI in ICU patients to identify and prevent the condition deterioration over time. We used the current kidney disease improving global outcomes (KDIGO) definition of AKI [9], defined as a serum creatinine level increase of 1.5 times baseline or 0.3 mg/dL within 48 hours. Given the poor specificity of the AKI classification [10] and inadequate data within our dataset, we opted to exclude urine output criteria. The baseline values were established as the initial serum creatinine measurement value taken within 24 hours after ICU admission. Therefore, we included only those 15 years older with at least two Scr values measurements.

As AKI events are associated with time, these bring barriers to data extraction and usage. However, in the MIMIC-III database, some features for time-related measurement characteristics were not defined, and the percentage of missing values ranges up to more than 60%. For example, time series, related to mean blood pressure, heart rate, temperature, hematocrit, albumin, etc., had between 55 and 90% of lost data. In situations of severe data incompleteness, imputing missing values can significantly diminish the overall sample's informational value and disrupt

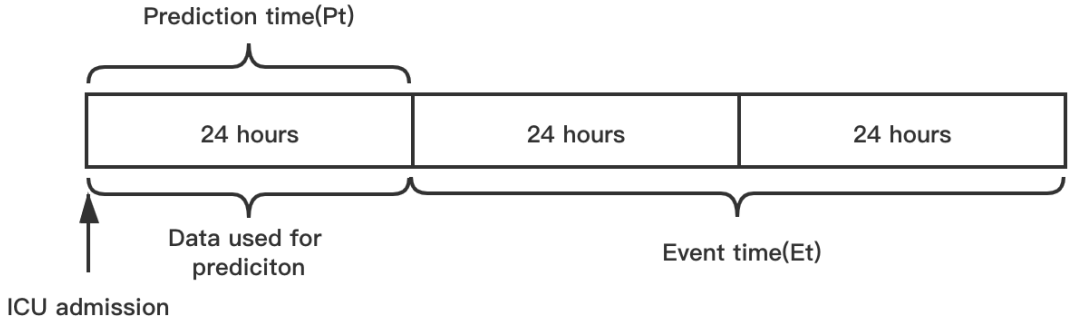

Fig. 2. Determining the appropriate time interval for data extraction. It is noted that the time window for the definition of acute kidney injury (AKI) is 24 hours after intensive care unit (ICU) admission.

the data's randomness. Therefore, we decided to exclude variables potentially related to kidney disease but had over 50% missing values. As a result, our foundational dataset included 15 characteristic variables.

Analytical Framework

We have proposed a framework for AKI prediction and interpretation to formulate a user-friendly and practical prediction model. Fig. 1 illustrates upstream data preprocessing and feature selection, downstream model construction and optimization. We utilized the criteria outlined in the sections discussing primary outcomes and predictive factors to construct our data set. Compared with other models, we implemented generative adversarial imputation networks (Gain) algorithms to address the issue of missing data. Finally, because the value range of the original data in the dataset used in this article has been determined, we chose the Min-max normalization method to normalize our data feature values to the range of [0,1], in order to preserve the relationship among the original data values.

In supervised machine learning (ML), models without an understanding of causality attempt to map any incorporated feature in their dataset to the target variable, even without a causal relationship. This can result in models that are inaccurate and prone to errors. Specifically, an excess of features can lead to overfitting due to increased model complexity. This study focuses on feature simplification for enhanced practicability within the framework. Feature selection involves choosing relevant features and removing redundant ones for a given classification task [11]. In this study, we employ popular feature selection methods to identify the most suitable feature subset for our AKI prediction model. This process simplifies the model, improves data compatibility with learning model classes, and facilitates interpretation by researchers.

Primary Outcome and Predictive Factors

The primary outcome of our study was the incidence of AKI within 3 days following ICU admission. We used the maximum creatinine values on days 2 and 3 to assess whether patients had AKI. The prediction time was shown with prediction time (Pt), and the AKI incidence time was 48 hours after Pt was shown with event time (Et). The schematic illustration of the time windows is presented in Fig. 2. The predicted result in Et is encoded as a binary categorical dependent variable with two categories: 0 (patient does not develop AKI at Et) and 1 (patient develops AKI at Et). As the ratio of 0 to 1 was close to 2:1, we used the synthetic minority oversampling technique (SMOTE) developed by Chawla *et al*. [12] to balance the data.

We endeavored to predict AKI onset by analyzing clinical data from the MIMIC III database. This data encompassed factors like patient age and gender, vital signs from the first day of ICU admission, laboratory values, and the total urine volume over 24 hours, among other details. The choice of candidate variables was guided by the application of well-established ICU disease risk prediction scoring systems (like acute physiology and chronic health evaluation (APACHE) and simplified acute physiology score (SAPS)), expert advice, and pertinent data within the dataset.

Prediction Model and Evaluation

In the present study, we utilized the extreme gradient boosting (XGBoost) algorithm to assess the risk of AKI among patients in the ICU. Among the machine learning methods currently in use, XGBoost represents an efficient embodiment of gradient-boosted decision trees. It has demonstrated outstanding performance across various models [13,14], exhibiting superior prediction accuracy and expedited processing time while maintaining computational economy and simplicity [15,16]. These features are particularly advantageous when navigating the sparse, high-dimensional clinical data inherent in AKI prediction. The unmatched accuracy and speed of XGBoost have solidified its prominence within the realm of applied machine learning.

Furthermore, we sought to evaluate the relative performance of XGBoost in comparison with six other prevalent machine learning methodologies, namely logistic re-

Table 1. Characteristics of patients in the medical information mart for intensive care (MIMIC-III) database.

Variable	Normal group (n = 10,316)	AKI group (n = 4902)
Age, years (mean ± SD)	64.63 ± 16.74	67.97 ± 14.97
Male, n (%)	5757 (55.81)	2907 (59.30)
GCS (15-GCS)	11.22 (14.00)	9.11 (10.00)
Baseline serum creatinine (mg/dL)	1.12 (0.90)	1.80 (1.30)
first_seq (mg/dL)	7.10 (6.00)	5.10 (4.00)
Frequency of urine test (/24 hours)	1.50 (1.00)	2.37 (2.00)
R (/minute)	18.97 (18.39)	19.23 (18.59)
PO_2 (mmHg)	127.81 (124.33)	126.86 (125.00)
WBC ($\times 10^9$/L)	11.58 (10.60)	12.82 (11.60)
BUN (mg/dL)	22.27 (17.00)	33.67 (26.33)
Na (mmol/L)	138.55 (138.75)	138.09 (138.00)
K (mmol/L)	4.07 (4.03)	4.31 (4.25)
PH	7.39 (7.39)	7.36 (7.37)
Glu (mg/dL)	135.68 (127.00)	143.79 (132.16)
TVU (mL/24 hr)	2044.71 (1790.00)	1497.04 (1225.00)

AKI, acute kidney injury; GCS, glasgow coma scale; R, respiratory rate; PO_2, partial pressure of oxygen; WBC, white blood cell; BUN, blood urea nitrogen; Na, sodium; K, potassium; PH, the potential of hydrogen; Glu, serum glucose; TVU, total volume urine (in 24 hours). All the clinical variables are represented as mean (median).

gression (LR), support vector machine (SVM), random forest classifiers (RF), artificial neural network (ANN), Gradient Boosting, and light gradient boosting machine (LightGBM). As highlighted by the study conducted by Chen *et al*. [15], the performance of the XGBoost model can be optimized by carefully adjusting the parameters. To facilitate this process, we employed GridSearchCV to automate the tuning of hyperparameters.

To develop and compare these models, we randomly partitioned the analytic cohort into a training set (constituting 70% of the cohort) and a test set (the remaining 30%). The risk prediction models for AKI were formulated using data from the training set. Throughout the model development phase, we segmented the training set into 5 folds for cross-validation purposes: 4 folds were allocated to model training. In contrast, the remaining fold was exclusively utilized for model validation. The average performance measure value derived from these 5 rounds was used to signify the performance of each model. The primary metric adopted for model selection and final reporting was the area under curve (AUC). The AUC is determined by calculating the area beneath the receiver operating characteristic (ROC) curve, which plots the false positive rate (FPR, 1-Specificity) against true positive rate (TPR, sensitivity) on the x and y-axes respectively, and is interpreted as a balance between specificity and sensitivity [17].

Statistical Analysis

We delineated the characteristics of the two groups: those with AKI and those without. The continuous variables were depicted as the mean ± standard deviation, while categorical data were denoted as a frequency (%). For non-normally distributed continuous data, we reported the median value. ROC curves were used for the comparison of prediction models. We used AUC and accuracy indices to show the diagnostic accuracy. As described above, a serious of models, including XGBoost, LR, SVM, RF, ANN, and LightGBM, were performed for prediction and comparison. All analyses in this study were carried out on the Jupyter Notebooks platform, using open-source libraries (scikit-learn, XGBoost) in Python 3.7 (the Python Software Foundation (PSF), Wilmington, Delaware, United States).

Results

Characteristics of AKI Patients

This study used MIMIC III for data extraction and model validation. Table 1 shows the characteristics of the study patients. The original cohort contains 15,218 patients, among which 4902 patients have developed AKI, accounting for about 32.21%. This analytic cohort's mean age was 65.70 ± 16.27 years, and 5757 (55.81%) were male. Compared with patients without AKI, patients with AKI were older (67.97 ± 14.97 versus 64.63 ± 16.74, p-value < 0.01) and more male (59.30% versus 55.81%, p-value < 0.01) (Table 1). Patients with AKI also had poor baseline kidney function and poorer glasgow coma scale (GCS), and were more likely to have diabetes.

Feature Selection

In this study, we utilized the recursive feature elimination with cross-validation (RFECV) technique to ascertain our model's optimal number of features. The RFECV method allows for a visual representation of the score corresponding to each feature subset, thus illustrating the pattern of feature elimination. Our findings suggested that subsets

Table 2. The performance comparison of the seven feature selection methods in extreme gradient boosting (XGboost) model test.

Categories	Methods	Features	8-features version		5-features version	
			Accuracy	AUC	Accuracy	AUC
Filter	F-test	frequency, BUN, Scr_baseline, K, GCS, first_seq, TVU, PH	0.780	0.826	0.767	0.811
	Pearson	frequency, Scr_baseline, BUN, K, GCS, first_seq, TVU, PH	0.780	0.826	0.767	0.811
	Chi-square test	frequency, GCS, first_seq, BUN, Scr_baseline, K, Age, TVU	0.770	0.820	0.764	0.808
Wrapper	Recursive feature elimination	frequency, Scr_baseline, R, PO_2, Age, TVU, K, PH	0.770	0.821	0.736	0.781
Embedded	L1-based feature selection	frequency, TVU, Scr_baseline, BUN, K, PO_2, GCS, PH	**0.781**	**0.835**	**0.769**	**0.814**
	RandomForest-based feature selection	frequency, Scr_baseline, TVU, PH, Age, PO_2, R, BUN	0.773	0.824	0.757	0.808
	Tree-based feature selection	frequency, Scr_baseline, BUN, TVU, K, GCS, PH, first_seq	0.780	0.826	**0.769**	**0.814**

The results show the mean values of the 5-fold cross-validation of each model. Features in the table are arranged according to the feature importance obtained by each algorithm. The best performance at each row is shown in bold. AUC, area under curve.

with 8 and 5 features were optimal. Consequently, we selected these feature subsets using various selection methodologies and subsequently determined the superior subset according to the cross-validation outcomes from the XGBoost models.

To appraise the efficacy of our feature selection algorithms, we utilized two widely-used metrics, namely accuracy and AUC. According to the results delineated in Table 2, the dataset constituted by the 8-feature subset slightly surpassed the one formed by the 5-feature subset in performance. The L1-based feature selection (least absolute shrinkage and selection operator (Lasso)) algorithm, significantly outshone the other six methods for both feature subsets. Intriguingly, despite eliminating nearly half the features, the model's performance did not markedly decline (accuracy: 0.786 versus 0.781; AUC: 0.844 versus 0.835). This observation suggests that the Lasso method can effectively streamline our model.

Model Development and Evaluation of Performance

Following feature selection, we fine-tuned the model in accordance with the current dataset. To gauge its generalization capability, we compared the experimental outcomes from the training data against those from the testing data. In this work, we mainly compared the performance of XGBoost with six other popular machine-learning techniques. Fig. 3 shows that the XGBoost model performs best in both data sets, and its generalization ability was better than other models. Overall, after the parameter tuning, the final AUC performance of XGBoost model indicators achieved are as follows: 0.849 (8 features) and 0.830 (5 features), higher than the result before parameter tuning (AUC = 0.835). This result further demonstrates the remarkable performance of the XGBoost model and suggests that it can be used to model AKI risk detection.

Comparison with Publicly Prediction Performance by a Systematic Review

We initiated a comprehensive literature review of studies centered around AKI prediction, utilizing 'MIMIC III' and 'AKI' as search keywords. This review encompassed studies published in PubMed between the years 2018 and 2023. Following an exhaustive screening of titles, subjects, abstracts, and full texts, we identified 79 studies for inclusion. However, as the objective of our research is to predict whether ICU patients will develop AKI, we excluded the following studies: those elucidating the correlation between indicators and AKI outcomes (n = 42), those forecasting AKI incidence and mortality risk in specific populations (n = 18); studies concentrating on the impact of medications on AKI (n = 7); and those predicting mortality risk in patients already diagnosed with AKI (n = 5) (see Appendix Fig. 4 for details). Consequently, we focused on the remaining 7 articles that predicted the risk of AKI in critically ill adult patients, summarized in Table 3 (Ref. [18–24]). As depicted in Table 3, the research objectives and prediction timelines of the first three articles align perfectly with our study. The prediction timelines for the following three articles were calculated based on the onset time of AKI, whereas the final article predicted AKI from the time of ICU admission.

Discussion

While literature is abundant on diverse prediction tasks concerning AKI [25], many of these studies use precise predictors or employ relatively small sample sizes

Table 3. Comparison with the state of the arts (for AKI prediction only) on the dataset MIMIC-III.

Studies	Purpose	Prediction time interval	Features	Sample size	Algorithms	AUC	Year of publication
Zimmerman *et al*. [18]	predict AKI in the first 24–72 h of ICU admission	24 h after admitted to ICU	32	23,950	Logistic Regression	0.78	2019
Li *et al*. [19]	predict AKI in the first 24–72 h of ICU admission	24 h after admitted to ICU	313 word features	16,560	Knowledge-guided CNN	0.779	2018
Sun *et al*. [20]	predict AKI in the first 24–72 h of ICU admission	24 h after admitted to ICU	22 structured features and 313 word features	16,558	Mixed-feature CNN	0.83	2019
Wang *et al*. [21]	predict the occurrence of AKI in next 24 and 48 h	24, or 48 h prior to AKI identification	38 features and 3235 distinct drugs	52,152	ETSM	0.95 (predicting AKI before 48 h)	2020
Le *et al*. [22]	AKI prediction at a 48-hour before onset	48h prior to AKI identification	7 structured features and clinical text data	12,347	CNN	0.86 (predicting AKI before 48 h)	2021
Gao *et al*. [23]	predict the occurrence of AKI in next 24, 48, and 72 h	24, 48, or 72 h prior to AKI identification	110 and electronic medical record data	30,020	ensemble model	0.895 (predicting AKI before 72 h)	2021
Qian *et al*. [24]	predict AKI in the first 72 h of ICU admission	After admission to ICU	17	17,205	LightGBM	0.905	2021
Proposed	predict AKI in the first 24–72 h of ICU admission	24 h after admitted to ICU	8	15,218	Gain-Lasso-XGBoost	0.849	
Proposed*	predict AKI in the first 24–72 h of ICU admission	24 h after admitted to ICU	5	15,218	Gain-Lasso-XGBoost	0.826	

*The data set used for model building contains 5 feature variables. ICU, intensive care unit; CNN, convolutional neural networks; ETSM, ensemble time series model; LightGBM, light gradient boosting machine; Gain-Lasso-XGBoost, generative adversarial imputation networks-least absolute shrinkage and selection operator-extreme gradient boosting.

DISCOVERY MEDICINE 2005 DM

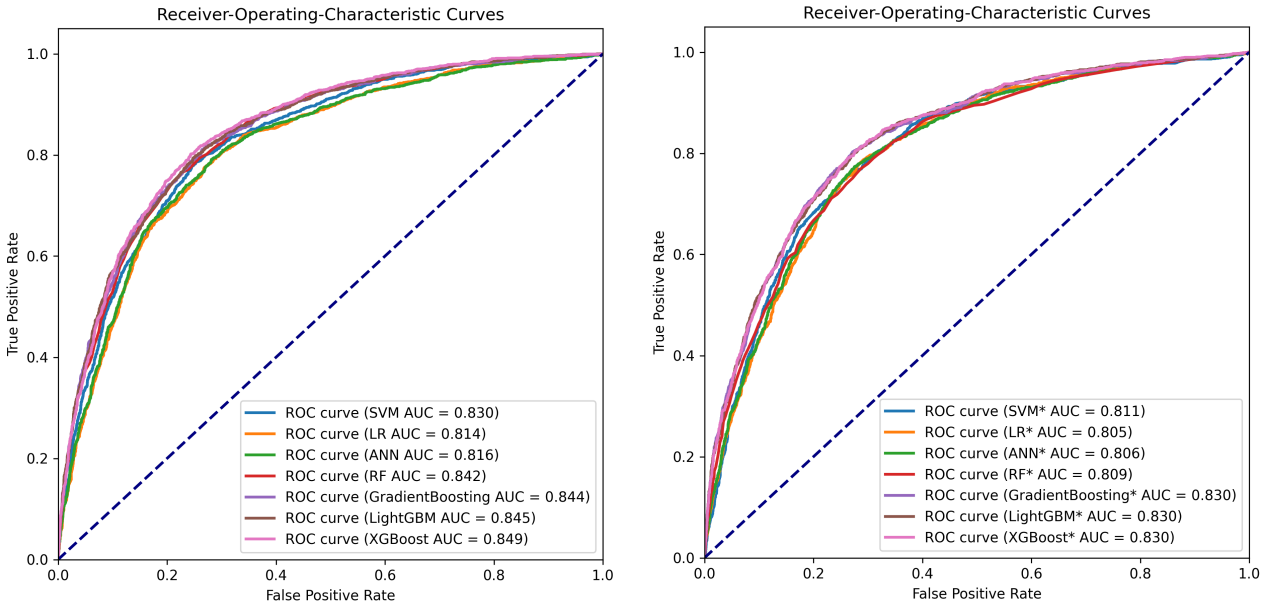

Fig. 3. Summarizes the area under curve (AUC) values for all models devised in this study, with the results determined using the test set. *The dataset utilized for model construction incorporates 5 feature variables.

[6,26,27]. These approaches must capture the high heterogeneity of the ICU patient population. Moreover, several models show only moderate performance when it comes to AUC or require an extensive set of variables [28,29], thus limiting their practical applications. For instance, Kate *et al.* [2] collected more than 30 features, including vital signs, lab results, comorbidities, medications, and so forth, to build a continuous AKI prediction model, yielding an AUC of 0.724. Contrastingly, our model outperformed previous studies [18–20] utilizing the MIMIC-III database and the same time window and achieving superior AUC scores (Gong *et al.* [30]: 0.781; Zimmerman *et al.* [18]: 0.783; Li *et al.* [19]: 0.779, respectively). The proposed Gain-Lasso-XGBoost framework in our study is more straightforward with easier-to-obtain features and demonstrates better generalization capability. Impressively, it yielded highly stable AUC results on two different datasets (comprising 8 and 5 features, respectively), underscoring its reliable predictive capacity for AKI. These results surpass those of the baseline model reported in the existing literature [18], affirming the effectiveness of our approach.

AKI in critical illness has complex pathophysiology that is difficult to represent, yet, artificial intelligence models can accurately represent it using routinely collected clinical data. Peer-reviewed literature on this topic is expanding rapidly, but most techniques are increasingly applied to large-scale, high-dimensional, multimodal ICU data, such as electronic medical records [31]. These studies present a diverse array of evidence, some of the high quality with the potential to enhance care and patient outcomes, while others of variable quality can potentially mislead or perplex investigators and clinicians. In our work, we have underscored the need for feature simplification within the framework for improved practicality. During feature selection, we also factored in the information about feature importance delivered by each feature subset, in conjunction with verifying the generalizability of features chosen by the feature screening method. According to the features results in Table 2, the baseline values of Scr (Scr_baseline), the number of Scr measurements within 24 hours of admission to ICU (frequency), and total volume urine (TVU) were included in all feature subsets, demonstrating that they play a key role in determining AKI risk prediction, consistent with the known pathophysiology of AKI [32]. In addition, blood urea nitrogen (BUN), potassium (K), and the potential of hydrogen (PH) appeared six times in seven feature subsets, also known as risk factors of AKI, and might represent patients with higher severity of illness. We aim for our identified key variables to garner attention and guide clinicians toward quicker and more efficient prediction of AKI risk in ICU patients. Furthermore, we hope our findings can help refine the focus of clinical validation and testing procedures. More importantly, all the variables mentioned above can be easily accessed at IUC, which will significantly improve the usefulness of our model.

The results of the system review also confirm that our proposed Gain-Lasso-XGBoost framework is simple and effective. Among all the models, we used the fewest features yet still demonstrated the best performance in the same type of study. While the other four studies had higher AUC and longer prediction time windows, our model was more suitable for ICU doctors to assess the patient's condition accurately and promptly take corresponding measures after admission.

Our study exhibits numerous strengths. Among them, significant work has been invested in feature processing, aiming to streamline the model and enhance its performance. Building upon these efforts, we introduced the Gain-Lasso-XGBoost framework. This framework not only simplifies the model and enhances its interpretability, more crucially, but significantly improves its applicability in a clinical setting. Secondly, the database used in this study is a publicly available and has large-scale ICU population size. Thirdly, the predictor variables we used can be available in the clinic practice, and this ensures that the values of these variables can be acquired in the clinical decision. Despite its strengths, this retrospective study does have some limitations. Firstly, the data used were sourced from a single-center database, which implies that further external validation and prospective trials are necessary to confirm the generalizability and evaluate the clinical utility of our findings. Secondly, due to the lack of pre-ICU admission serum creatinine data, in line with previous studies, we used the lowest serum creatinine level recorded after ICU admission as the baseline.

Conclusions

In light of the significant health implications and additional resource demands associated with AKI, we've devised a Gain-Lasso-XGBoost framework. This framework aims to facilitate early detection of AKI in critically ill adults by harnessing data from the initial 24 hours following ICU admission. By aiding ICU clinicians in swiftly identifying and addressing high-risk AKI patients, this approach can play a vital role in enhancing patient prognosis, a key goal in AKI management.

Availability of Data and Materials

The datasets supporting the conclusions of this article are available in the https://mimic.physionet.org/.

Author Contributions

YL, HLB and HL designed the research study. YL and WLY contributed to the conduct of the study. YWS extracted and processed the data from MIMIC-III. YL and HLB conduced the model development and data analysis. The results were analyzed, interpreted and discussed by YL, HL and WLY. YL drafted the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The dataset supporting the conclusions of this article is available in the Medical Information Mart for Intensive Care version III (MIMIC-III) version 1.4. The database is public de-identified databases thus informed consent and approval of the Institutional Review Board was waived. Our access to the database was approved after completion of the Collaborative Institutional Training Initiative (CITI program) web-based training course called "Data or Specimens Only research" (Record ID: 29889121). More details are available at https://mimic.physionet.org/gettingstarted/access/#request-access-to-mimic-iii.

Acknowledgment

Not applicable.

Funding

This work was supported by CAMS Innovation Fund for Medical Sciences (grant number 2022-I2M-2-001).

Conflict of Interest

The authors declare no conflict of interest.

Appendix

See Fig. 4.

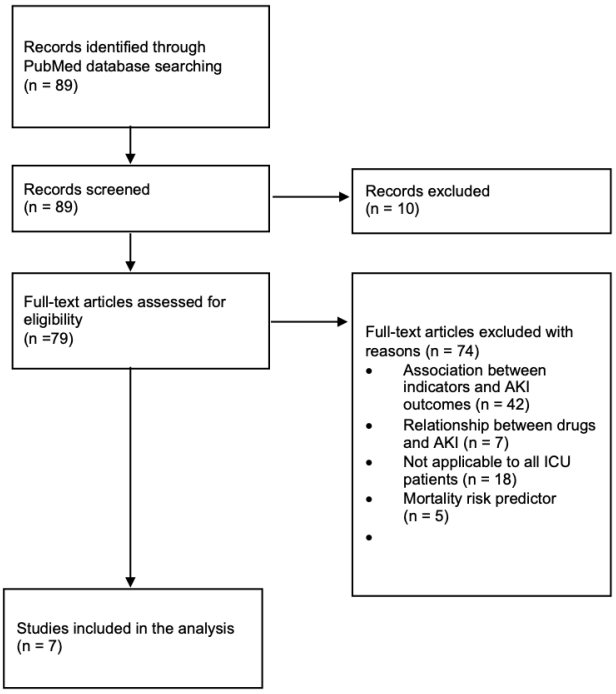

Fig. 4. PRISMA flow diagram of the study selection.

References

[1] Wang B, Ding S, Liu X, Li X, Li G. Predictive classification of ICU readmission using weight decay random forest. Future Generation Computer Systems. 2021; 124: 351–360.

[2] Kate RJ, Pearce N, Mazumdar D, Nilakantan V. A continual prediction model for inpatient acute kidney injury. Computers in Biology and Medicine. 2020; 116: 103580.

[3] Hoste EAJ, Kellum JA, Selby NM, Zarbock A, Palevsky PM, Bagshaw SM, *et al*. Global epidemiology and outcomes of acute kidney injury. Nature Reviews. Nephrology. 2018; 14: 607–625.

[4] Bhosale SJ, Kulkarni AP. Biomarkers in Acute Kidney Injury. Indian Journal of Critical Care Medicine. 2020; 24: S90–S93.

[5] Flechet M, Falini S, Bonetti C, Güiza F, Schetz M, Van den Berghe G, *et al*. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. Critical Care. 2019; 23: 282.

[6] Rashidi HH, Sen S, Palmieri TL, Blackmon T, Wajda J, Tran NK. Early Recognition of Burn- and Trauma-Related Acute Kidney Injury: A Pilot Comparison of Machine Learning Techniques. Scientific Reports. 2020; 10: 205.

[7] Cui H, Shu S, Li Y, Yan X, Chen X, Chen Z, *et al*. Plasma Metabolites-Based Prediction in Cardiac Surgery-Associated Acute Kidney Injury. Journal of the American Heart Association. 2021; 10: e021825.

[8] Johnson AEW, Pollard TJ, Shen L, Lehman LWH, Feng M, Ghassemi M, *et al*. MIMIC-III, a freely accessible critical care database. Scientific Data. 2016; 3: 160035.

[9] Kellum J A, Lameire N, Aspelin P, Barsoum RS, Burdmann EA, Goldstein SL, *et al*. Kidney disease: improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. Kidney International Supplements. 2012; 2: 1–138.

[10] Thau MR, Bhatraju PK. Sub-Phenotypes of Acute Kidney Injury: Do We Have Progress for Personalizing Care? Nephron. 2020; 144: 677–679.

[11] Zhou P, Li P, Zhao S, Wu X. Feature Interaction for Streaming Feature Selection. IEEE Transactions on Neural Networks and Learning Systems. 2021; 32: 4691–4702.

[12] Chawla NV, Bowyer KW, Hall L, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. 2002; 16: 321–357.

[13] Karbassiyazdi E, Fattahi F, Yousefi N, Tahmassebi A, Taromi AA, Manzari JZ, *et al*. XGBoost model as an efficient machine learning approach for PFAS removal: Effects of material characteristics and operation conditions. Environmental Research. 2022; 215: 114286.

[14] Xia Y, Li X, Chen X, Lu C, Yu X. Inferring Retinal Degeneration-Related Genes Based on Xgboost. Frontiers in Molecular Biosciences. 2022; 9: 843150.

[15] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. 2016; 785–794.

[16] Mousa S R, Bakhit P R, Ishak S. An extreme gradient boosting method for identifying the factors contributing to crash/near-crash events: a naturalistic driving study. Canadian Journal of Civil Engineering. 2019; 46: 712–721.

[17] Ferri C, Hernández-Orallo J, Flach PA. A coherent interpretation of AUC as a measure of aggregated classification performance. in Proceedings of the 28th International Conference on Machine Learning (ICML-11). 2011; 657–664.

[18] Zimmerman LP, Reyfman PA, Smith ADR, Zeng Z, Kho A, Sanchez-Pinto LN, *et al*. Early prediction of acute kidney injury following ICU admission using a multivariate panel of physiological measurements. BMC Medical Informatics and Decision Making. 2019; 19: 16.

[19] Li Y, Yao L, Mao C, Srivastava A, Jiang X, Luo Y. Early Prediction of Acute Kidney Injury in Critical Care Setting Using Clinical Notes. Proceedings. IEEE International Conference on Bioinformatics and Biomedicine. 2018; 2018: 683–686.

[20] Sun M, Baron J, Dighe A, Szolovits P, Wunderink RG, Isakova T, *et al*. Early Prediction of Acute Kidney Injury in Critical Care Setting Using Clinical Notes and Structured Multivariate Physiological Measurements. Studies in Health Technology and Informatics. 2019; 264: 368–372.

[21] Wang Y, Wei Y, Yang H, Li J, Zhou Y, Wu Q. Utilizing imbalanced electronic health records to predict acute kidney injury by ensemble learning and time series model. BMC Medical Informatics and Decision Making. 2020; 20: 238.

[22] Le S, Allen A, Calvert J, Palevsky PM, Braden G, Patel S, *et al*. Convolutional Neural Network Model for Intensive Care Unit Acute Kidney Injury Prediction. Kidney International Reports. 2021; 6: 1289–1298.

[23] Gao W, Wang J, Zhou L, Luo Q, Lao Y, Lyu H, *et al*. Prediction of acute kidney injury in ICU with gradient boosting decision tree algorithms. Computers in Biology and Medicine. 2021; 140: 105097.

[24] Qian Q, Wu J, Wang J, Sun H, Yang L. Prediction Models for AKI in ICU: A Comparative Study. International Journal of General Medicine. 2021; 14: 623–632.

[25] Han H, Wang WY, Mao BH. 'Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning,' Advances in Intelligent Computing: International Conference on Intelligent Computing ICIC 2005. Hefei, China, 23–26 August 2005, Proceedings, Part I 1. Springer: Berlin Heidelberg. 2005; 878–887.

[26] Wong B, St Onge J, Korkola S, Prasad B. Validating a scoring tool to predict acute kidney injury (AKI) following cardiac surgery. Canadian Journal of Kidney Health and Disease. 2015; 2: 3.

[27] Lei VJ, Luong T, Shan E, Chen X, Neuman MD, Eneanya ND, *et al*. Risk Stratification for Postoperative Acute Kidney Injury in Major Noncardiac Surgery Using Preoperative and Intraoperative Data. JAMA Network Open. 2019; 2: e1916921.

[28] Zhao X, Lu Y, Li S, Guo F, Xue H, Jiang L, *et al*. Predicting renal function recovery and short-term reversibility among acute kidney injury patients in the ICU: comparison of machine learning methods and conventional regression. Renal Failure. 2022; 44: 1326–1337.

[29] Zhang Y, Zhang X, Xi X, Dong W, Zhao Z, Chen S. Development and validation of AKI prediction model in postoperative critically ill patients: a multicenter cohort study. American Journal of Translational Research. 2022; 14: 5883–5895.

[30] Gong K, Lee HK, Yu K, Xie X, Li J. A prediction and interpretation framework of acute kidney injury in critical care. Journal of biomedical informatics. 2021; 113: 103653.

[31] Ozrazgat-Baslanti T, Loftus TJ, Ren Y, Ruppert MM, Bihorac A. Advances in artificial intelligence and deep learning systems in ICU-related acute kidney injury. Current Opinion in Critical Care. 2021; 27: 560–572.

[32] Basile DP, Anderson MD, Sutton TA. Pathophysiology of acute kidney injury. Comprehensive Physiology. 2012; 2: 1303–1353.