Machine Learning Model-Based Early Detection of Acute Kidney Injury in Hypertensive Patients

Lachlan Mcalpine, 1475293
Meghana Ganapa, 1436378
Rongcong Lin, 1520403
Yueyang Wu, 1345511

Abstract

Objectives: Acute kidney injury (AKI) is a common and serious combination among hypertensive patients. Early detection and timely intervention may improve patients outcomes and alleviate healthcare burdens. This study aimed to develop and validate machine learning (ML) models for the early prediction of AKI in intensive care unit (ICU) patients with hypertension.

Materials and Methods: Hypertensive patients were retrieved from the Medical Information Mart for Intensive Care (MIMIC-IV) database. Six machine learning methods were implemented based on their significant demographic and clinical information, and the optimal model was selected based on the area under the receiver operating characteristic (AUROC) and calibration plot.

Results: In total, 12,066 hypertensive patients were included in the study with a median age of 67 years and 5,251 (43.5%) with AKI. A set of 24 clinical features was selected for model development. Both the eXtreme Gradient Boosting (XGBoost) model and Logistic regression (LR) demonstrated favorable discrimination power achieving AUROC of 0.825 and 0.818 respectively. Additionally the calibration plot of the model revealed close alignment between the prediction probabilities and the observed outcomes.

Conclusion: The findings of this study indicate that a machine learning-driven AKI prediction tool has acceptable performance in the early detection of AKI in hypertensive patients. However, a larger model that includes comprehensive features, multi-center collaboration, and clinical validation is necessary before it can benefit society.

Introduction

Hypertension is a leading contributor to the global burden of disease and remains the major modifiable risk factor for cardiovascular conditions and kidney damage[1]. Moreover, the kidney is a vulnerable organ affected by hypertension-induced injury[2], with hypertension being linked to an increased risk of AKI in several clinical settings.[3]

In clinical practice, the diagnosis of AKI is primarily based on changes in serum creatinine (SCr) or urine output (UO).[4] However, acute changes in creatinine often significantly lag behind the onset of renal injury,[5] and currently, there are no effective treatments to reverse severe AKI once it has developed.[6] Therefore, early detection and timely intervention of high-risk patients are critical to ensure the appropriate allocation of limited clinical resources, which may ultimately improve patient outcomes.[7,8]

Prediction of AKI has always been a hot topic in healthcare research. For instance, using interpretable machine learning techniques to predict the mortality within the hospital for critically ill AKI patients, the model showed a favorable prediction accuracy.[9] Moreover, many researchers have refined predictive models for specific patient populations, such as diabetic patients,[10] sepsis patients[11] and post-cardiac surgery patients.[12] These experimental approaches provide valuable insights for future studies and may contribute to clinical validation.

However, there have been few attempts to predict AKI early in hypertensive patients, and limited analysis of the factors influencing their outcomes. In this study, we aimed to apply machine learning approaches on an electronic health record data (EHR) - MIMIC-IV to predict the presence of AKI in critically ill patients with hypertension and to identify significant factors influencing this prognosis.

Methods

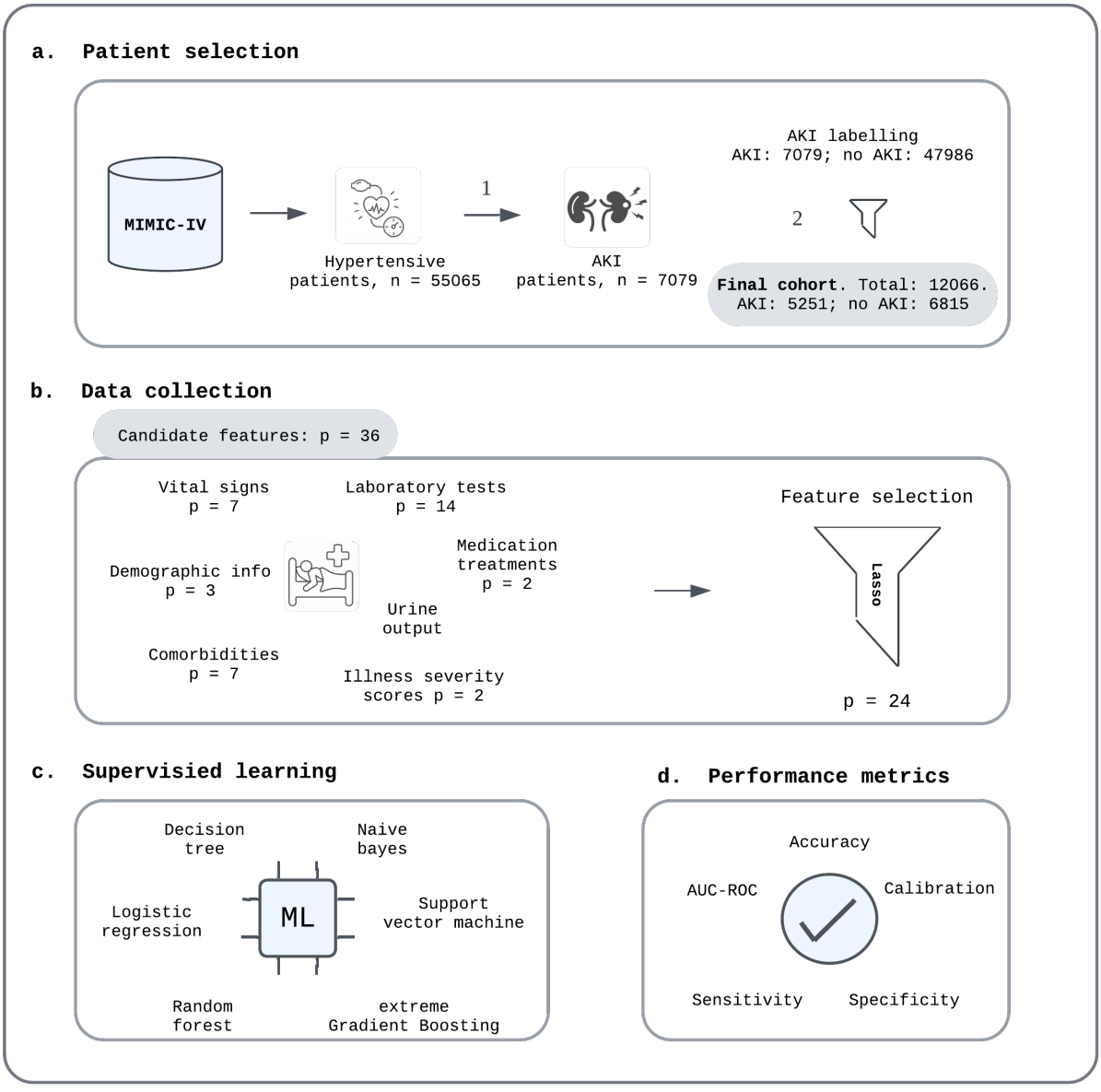
Ethics statement

The data used in this study was sourced from the publicly available MIMIC-IV database, which has been fully de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPPA) Privacy Rule. Since studies conducted using de-identified data are considered nonhuman subject research, no ethical approvals were required for this study. And there are no plans to redistribute the retrieved data.

Study population

Patients who received hypertensive medication during admission were treated as hypertensive patients. The hypertensive medications were based on those listed by the American Heart Association[13] website. We then assessed whether they developed AKI or not during their first admission based on the kidney disease: Improving Global Outcomes (KDIGO)[14] criteria as follows: increase in serum creatinine (SCr) $\geq$ 0.3 mg/dl within 48 h or urine volume $<$ 0.5 mL/kg/hour for 6 h. Many of these hypertensive patients lacked sufficient demographic and clinical information in the database. To ensure the data quality during ML modeling. We selected only those hypertensive patients who had Glasgow Coma Scale (GCS) records within first 24 h of ICU admission. This approach was inspired by the research of Hu et al[9] in 2022, which found GCS to be one of the most important variables contributing to the models for predicting the in-hospital mortality of AKI. Moreover, patients with missing >20% individual data at admission were excluded.

Figure 1: Overview of the study



1. AKI criteria: increase in SCr >= 0.3 mg/dl within 48 h; or urine volume < 0.5 mL/kg/hour for 6 h
2. Excluded patients without GCS records within first 24h of ICU admission; and patients with > 20% missing data (i.e., > 8 features)

Data collection and preprocessing

For each included patient, the following information was used in this study: (1) demographic features, including age, gender, and average weight; (2) comorbidities, including diabetes, liver disease, renal disease, cardiovascular disease, chronic pulmonary disease, tumor and acquired immune deficiency syndrome; (3) the mean values of vital signs during first 24h of admission, including heart rate, systolic blood pressure, diastolic blood pressure, mean arterial pressure, respiratory rate, body temperature, and oxygen saturation (SpO_2); (4) the mean values of laboratory tests within first 24h, including white blood cell, hematocrit, hemoglobin, platelets, blood urea nitrogen, international normalized ratio, anion gap, bicarbonate, blood glucose, serum creatinine, serum sodium, serum potassium, serum calcium, and serum chloride; (5) medications received during first 24h of admission, including analgesics and antibiotics; (6) the total urine volume within first 24h; (7) the mean values of illness severity scores during first 24h, including GCS and Simplified Acute Physiology Score II (SAPS-II).

In order to minimize the bias from missing data, a multiple imputation method with 5 iterations was used to deal with missing features since this procedure can better capture the uncertainty behind the missing value.[9,11] In addition, to reduce overfitting and improve model interpretability, Lasso Regression with 5-fold stratified cross validation was applied to shrink less important features.

Statistical analysis

The normality of continuous features were determined by Kolmogorov-Smirnov test. And they were presented as the mean ± standard deviation for normal distribution or median and interquartile range for abnormal distribution. In terms of the categorical features, they were all presented as the number and proportion. Moreover, Student's t test and Pearson's chi-square test were employed to compare differences in the clinical data between AKI and no AKI cohort. A P-value below 0.05 was considered as statistically significant.

Machine learning and validation method

To avoid the bias of algorithms assigning more importance to features with larger value, all variables were standardized based on their mean and standard deviation firstly. We then randomly split all patients into a training set (75%) and a validation set (25%). Lasso[15] Regression was used to find out important features as the input of the ML models, the output of which is a probability of AKI occurring during ICU admission.

The six ML methods: Logistic Regression[16] (LR), Decision Trees[17] (DT), Naive Bayes[18] (NB), Support Vector Machine[19] (SVM), Random Forest[20] (RF) and extreme Gradient Boosting[21] (XGBoost) were implemented to predict the in-hospital AKI occurrence among hypertensive patients. The hyperparameters of RF and XGBoost were optimized under 10-fold stratified grid search cross validation in the training dataset by maximizing the area under the receiver operating characteristic curve (AUROC).

The performance of the predictive models was mainly determined by the capacity of discrimination and calibration. The discrimination was quantitatively evaluated by AUROC. And the calibration was visually accessed through the graphical representations of the consistency of the predicted probability and observed outcomes. Then the model with favorable AUROC was selected as the optimal model, and SHAP[22] algorithm was used to visualize the significant features that influence the occurrence of AKI.

Results

Participants

There were 55065 patients received hypertensive medications in MIMIC-IV database, they are known as treated hypertensive patients.[23] Among this cohort, 7079 individuals experienced acute kidney injury during their first ICU admission. To construct a patient health profile with low missing data ratio, many patients were excluded according to the exclusion criteria in Figure 1. Finally, a total of 12066 patients were enrolled in this study, of which 5251 (43.5%) had AKI during ICU admission.

Table 1: Patients' characteristics (Part 1)

Characteristics	**Total** (n=12066)	**AKI** (n=5251)	**Non-AKI** (n=6815)	**P-value**
Demographics				
Age, years	67 (56–77)	68 (57–77)	67 (56–77)	0.056
Gender (Male)	6840 (56.7)	3213 (61.2)	3627 (53.2)	< 0.001
Weight, kg	78.9 (66.1–93.7)	81 (68.1–97.1)	77.6 (64.7–91)	< 0.001
Comorbidities				
Diabetes	3628 (30.1)	1851 (35.3)	1777 (26.1)	< 0.001
Liver disease	1820 (15.1)	1109 (21.1)	711 (10.4)	< 0.001
Renal disease	4614 (38.2)	3192 (60.8)	1422 (20.9)	< 0.001
Cardiovascular disease	5285 (43.8)	2671 (50.9)	2614 (38.4)	< 0.001
Chronic pulmonary disease	3827 (31.7)	1936 (36.9)	1891 (27.7)	< 0.001
Tumor	2918 (24.2)	1244 (23.7)	1674 (24.6)	0.276
Acquired immune deficiency syndrome	114 (0.9)	52 (1.0)	62 (0.9)	0.720

Table 2: Patients' characteristics (Part 2)

Vital signs on Day 1				
Heart rate, bpm	81.8 (72.6–93)	83.9 (74.5–96.2)	80.3 (71.3–90.6)	< 0.001
Systolic blood pressure, mmHg	119.1 (109.1–132.1)	115 (106.4–127.2)	122.6 (112–134.6)	< 0.001
Diastolic blood pressure, mmHg	64.2 (57.4–72.8)	62 (55.8–69.8)	65.9 (58.9–74.7)	< 0.001
Mean arterial pressure, mmHg	79.9 (73.4–88.4)	77.4 (71.7–85.5)	82.1 (75–90.6)	< 0.001
Respiratory rate	18.6 (16.6–21)	19.1 (16.8–22)	18.2 (16.4–20.4)	< 0.001
Body temperature, °C	36.8 (36.6–37.1)	36.8 (36.6–37.1)	36.8 (36.7–37)	0.002
SpO_2, %	97 (95.6–98.4)	97.2 (95.7–98.5)	96.8 (95.5–98.2)	< 0.001
Laboratory findings on Day 1				
White blood cell, K/uL	10.2 (7.8–13.3)	11 (8.4–14.7)	9.6 (7.5–12.3)	< 0.001
Hematocrit, %	35.7 (31.4–39.6)	34.3 (30.2–38.5)	36.6 (32.5–40.1)	< 0.001
Hemoglobin, g/dL	11.8 (10.3–13.1)	11.3 (9.8–12.7)	12.2 (10.7–13.4)	< 0.001
Platelets, K/uL	198.9 (154.3–251)	191 (145.1–242.9)	204 (161.8–255.1)	< 0.001
Blood urea nitrogen, mg/dL	18 (13.5–26)	22 (15.7–34.5)	16 (12.3–21)	< 0.001
International normalized ratio	1.2 (1.1–1.4)	1.2 (1.1–1.4)	1.1 (1–1.3)	< 0.001
Anion gap, mEq/L	14.8 (13–16.7)	15.2 (13.3–17.5)	14.5 (13–16)	< 0.001
Bicarbonate, mmol/L	23.7 (21.5–25.6)	23 (22–25.2)	24 (22.2–26)	< 0.001
Blood glucose, mg/dL	123.7 (105.3–153.5)	132 (110.3–167.1)	118.5 (102.5–144.2)	< 0.001
Serum creatinine, mg/dL	0.9 (0.8–1.3)	1.1 (0.8–1.7)	0.8 (0.7–1)	< 0.001
Serum sodium, mEq/L	139 (137–141)	139 (136.5–141)	139 (137–141)	0.020
Serum potassium, mEq/L	4.2 (3.9–4.5)	4.2 (4–4.6)	4.1 (3.9–4.4)	< 0.001
Serum calcium, mg/dL	8.6 (8.2–9)	8.5 (8–8.9)	8.7 (8.4–9.1)	< 0.001
Serum chloride, mg/dL	103 (100.2–105.5)	103 (99.8–106)	103 (100.5–105)	0.048
Medication treatments on Day 1				
Analgesics	6161 (51.1)	2931 (55.8)	3230 (47.4)	< 0.001
Antibiotics	6803 (56.4)	3604 (68.6)	3199 (46.9)	< 0.001
Urine output on Day 1, mL	1250 (750-1940)	1170 (664-1875)	1315 (805-1986)	0.002
Severity scores of illness on Day 1				
GCS	15 (14.6-15)	15 (14.6-15)	15 (14.6-15)	0.486
SAPS-II	33 (26-41)	39 (31-47)	29 (23-36)	< 0.001

All values were reported as median (IQR) or no. (%).
Abbreviations: GCS, Glasgow Coma Scale; SAPS-II, Simplified Acute Physiology Score II.

Patients baseline characteristics differed in several important ways between AKI and non-AKI groups, it were summarized in Table 1 and Table 2. Male hypertensive patients had a higher likelihood of developing AKI compared to female patients during their ICU stay. Additionally, patients who experienced AKI tended to have a higher body weight. The two most prevalent comorbidities among hypertensive AKI patients were existing renal disease and cardiovascular conditions, affecting nearly 60% and 50%, respectively.

Feature selection

An initial set of thirty-six candidate features were put into a Lasso regression model with 5-fold stratified cross validation. As a result, twenty-four features were identified as significant predictors for AKI, as illustrated in appendix Figure 1.

Model development and validation

The enrolled patients were randomly divided into a training cohort of 9049 individuals, of which 3929 (43.3%) developed AKI, and a validation cohort of 3017 individuals, with 1328 (44%) experiencing AKI. Twenty-four health characteristics of the patients were chosen as input features for six machine learning models. After 10-fold stratified cross validation, XGBoost achieved the highest AUROC of 0.825, with RF at 0.821, LR at 0.818, SVM at 0.816, NB at 0.788 and DT at 0.653, as shown in Figure 2A. Mo et al[10] suggest that an area under the ROC curve greater than 0.7 indicates good predictive performance. Based on this criterion, the decision tree may not be well-suited for binary classification tasks in predicting AKI occurrence. Nevertheless, this does not preclude the possibility that it may demonstrate sound predictive performance in identifying different stages of AKI.

To further evaluated model performance, precision, sensitivity, and specificity were assessed and summarized in Table 3. The overall performance of XGBoost was comparable to that of Logistic regression, random forest and Support vector machine across the four evaluation metrics. Their specificity ranged from 0.826 to 0.844, indicating that approximately 16% to 18% nonAKI pa-

tients might be incorrectly predicted as having AKI. And a less optimistic outcome was observed in their sensitivity, which ranged from 0.622 to 0.653, meaning that about 35% AKI patients would not be detected early by these models.

Additionally, calibration analysis offered insight into the reliability of a model's probability predictions. XGBoost and logistic regression demonstrated the closest alignment between predicted probabilities and the observed AKI frequency, as illustrated in Figure 2B. Therefore, based on both discrimination and calibration metrics, XGBoost and Logistic regression were identified as the acceptable models for predicting in-hospital AKI occurrence in the hypertensive cohort.

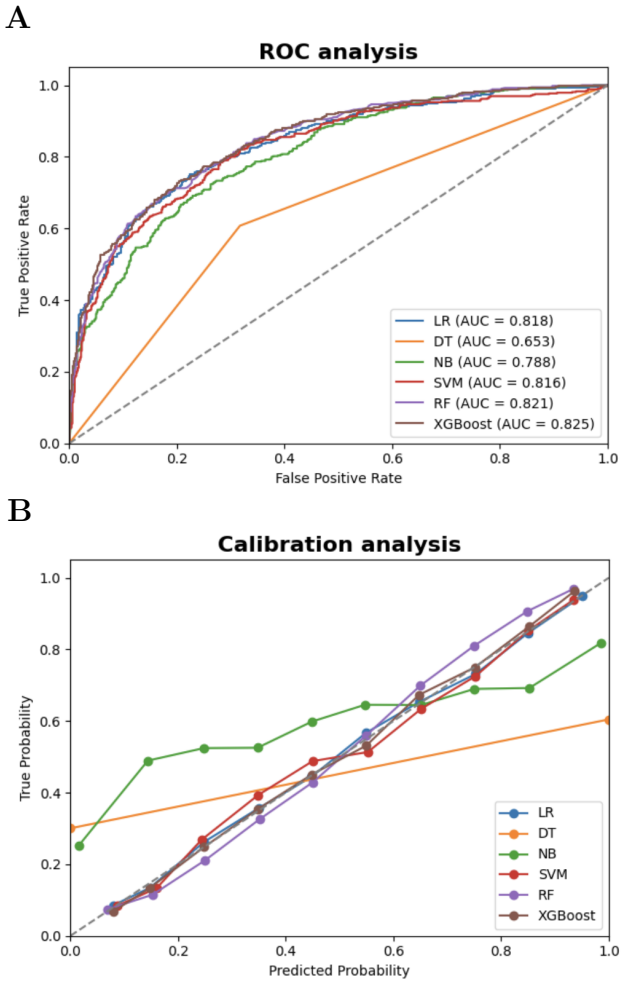


Figure 2: AUC-ROC and Calibration Curve Analysis of Model Predictions

Table 3: Model performance metrics

Model	Accuracy	Sensitivity	Specificity	AUROC
LR	0.751	0.645	0.832	0.818
DT	0.658	0.616	0.690	0.653
NB	0.716	0.494	0.887	0.788
SVM	0.747	0.622	0.844	0.816
RF	0.750	0.651	0.826	0.821
XGBoost	0.751	0.653	0.826	0.825

Abbreviations: LR: logistic regression; DT: decision tree; NB: naive bayes; SVM: support vector machine; RF: random forest; XGBoost: eXtreme gradient boosting.

Feature relevance

XGBoost was chosen to investigate feature importance in prediction using the shapley algorithm, as shown in Figure 3. The top four impactful factors contributing to the development of in-hospital AKI were identified as existing renal disease, SAPS-II, serum creatinine levels, age, and serum calcium.

Discussion

The prediction of AKI has two major approaches: one focuses on identifying novel biomarkers to detect early stage renal injury, while the other leverages machine learning models based on EHR to predict its occurrence. Previous studies have developed numerous models to predict in-hosptial AKI,[9,24,25] with some narrowing their focus to specific populations, achieving favorable performance. For instance, A Cox regression model was developed for diabetic patients[10] in China, achieving an AUROC of 0.774; An XGBoost model was developed using MIMIC-III data for sepsis patients,[11] obtaining an AUCROC of 0.821; and in Germany, an XGBoost model was implemented to focus on post-cardiac surgery patients,[26] achieving an AUCROC of 0.88.

To the best of our knowledge, this study is the first to target hypertensive patients as distinct cohort for AKI prediction. Our findings revealed that both the boosting classifier (XGBoost) and the linear model (logistic regression) demonstrated comparable performance in terms of discrimination and calibration. This contrasts slight with the prevailing consensus[9,11,26] that XGBoost generally outperforms logistic regression in AKI prediction.

Additionally, we introduced a novel data collection method by using Glasgow Coma Scale (GCS) score as an anchor point for selecting eligible patients, which helped minimize missing data, achieving a missing rate of less than 1%. This method reduces the uncertainty typically associated with imputation. A plausible explanation for the more complete data in patients with GCS scores is that these individuals often present with more severe conditions, warranting more frequent and thorough monitoring, thus resulting in more complete health records.

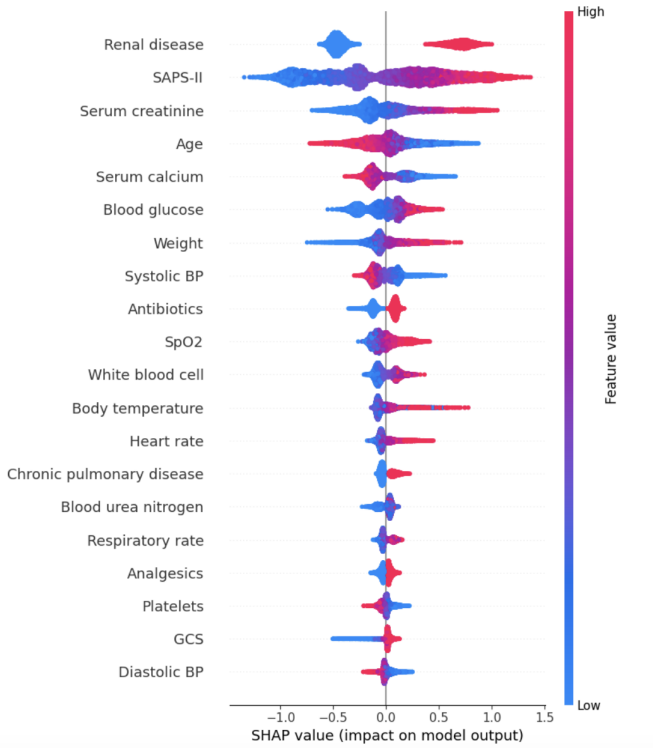
The SHAP algorithm was employed to interpret the XGBoost model and identify the most influential factors associated with AKI occurrence. Existing renal disease, higher SAPS-II score, elevated serum creatinine levels within the first 24 hours, elevated blood glucose, and obesity were identified as factors that may increase the likelihood of AKI during admission. Conversely, slightly high systolic blood pressure and higher serum calcium levels may reduce the risk of AKI.

However, this study had several limitations. First, we included only patients who had GCS records within the first 24 hours of admission, which inevitably narrowed the cohort to a more severely ill group, introducing selection bias. Second, only twenty-four clinical features were considered in model construction, some potential

risk factors may have been missed. Third, the MIMIC-IV dataset is derived from a single center in the United States, potentially limiting the generalization of our prediction model to other populations.

Therefore, future research with comprehensive feature collection (e.g., considering the change of certain indicators over 24h instead of static values, or including the lowest and highest values for better capturing dynamic patterns). Additionally, multi-center collaborations and clinical validation are necessary for building a prediction model that can benefit patients health and alleviate healthcare burdens.

Figure 3: Model interpretation



Conclusion

Machine learning models based on EHR were developed and validated with acceptable performance in the early prediction of AKI in hypertensive patients. Both XGBoost and logistic regression demonstrated favorable discrimination and calibration through internal validation.

Contribution

Author	Conceptualization	Methodology	Software	Validation	Formal Analysis	Investigation	Resources	Data Curation	Writing Original Draft	Writing Review & Editing	Visualization
Meghana G.	✓		✓	✓		✓	✓	✓	✓		
Yueyang W.	✓	✓	✓	✓		✓	✓		✓		✓
Lachlan M.	✓		✓	✓	✓		✓		✓		✓
Rongcong L.	✓		✓	✓	✓	✓	✓		✓	✓	

References

1. Hamrahian SM, Falkner B. Hypertension in chronic kidney disease. *Hypertension: from basic research to clinical practice* 2017; 956:307–25.
2. Griffin KA. Hypertensive kidney injury and the progression of chronic kidney disease. *Hypertension* 2017; 70(4):687–94.
3. James MT, Grams ME, Woodward M, et al. A meta-analysis of the association of estimated GFR, albuminuria, diabetes mellitus, and hypertension with acute kidney injury. *Am J Kidney Dis* 2015; 66(4):602–12.
4. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* 2012; 120(4):c179–84.
5. Doi K, Yuen PST, Eisner C, et al. Reduced production of creatinine limits its use as marker of kidney injury in sepsis. *J Am Soc Nephrol* 2009; 20(6):1217–21.
6. Zeng X, Shi S, Sun Y, et al. A time-aware attention model for prediction of acute kidney injury after pediatric cardiac surgery. *J Am Med Inform Assoc* 2022; 30(1):94–102.
7. Al-Jaghbeer M, Dealmeida D, Bilderback A, et al. Clinical decision support for in-hospital AKI. *J Am Soc Nephrol* 2018; 29(2):654–60.
8. Hodgson LE, Selby N, Huang TM, et al. The role of risk prediction models in prevention and management of AKI. *Semin Nephrol* 2019; 39(5):421–30.
9. Hu C, Tan Q, Zhang Q, et al. Application of interpretable machine learning for early prediction of prognosis in acute kidney injury. *Comput Struct Biotechnol J* 2022; 20:2861–70.
10. Mo M, Pan L, Huang Z, et al. Development and validation of a prediction model for survival in diabetic patients with acute kidney injury. *Front Endocrinol* 2021; 12:737996.
11. Yue S, Li S, Huang X, et al. Machine learning for the prediction of acute kidney injury in patients with sepsis. *J Transl Med* 2022; 20(1):215.
12. Ryan CT, Zeng Z, Chatterjee S, et al. Machine learning for dynamic and early prediction of acute kidney injury after cardiac surgery. *J Thorac Cardiovasc Surg* 2023; 166(6):e551–64.
13. Types of blood pressure medications [Internet].[Dallas (TX): American Heart Association]. [updated 2024 May 24; cited 2024 Oct 29]. Available from: `https://www.heart.org/en/health-topics/high-blood-pressure/changes-you-can-make-to-manage-high-blood-pressure/types-of-blood-pressure-medications`
14. Stevens PE, Levin A. Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group M. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med* 2013; 158(11):825–30.
15. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J Roy Stat Soc B Met* 1996; 58(1):267–88.
16. Pregibon D. Logistic Regression Diagnostics. *Ann Stat* 1981; 9(4):705–24.
17. Safavian SR, Landgrebe D. A Survey of Decision Tree Classifier Methodology. *leee T Syst Man Cyb* 1991; 21(3):660–74.

18. Zolnierek A, Rubacha B. The Empirical Study of the Naive Bayes Classifier in the Case of Markov Chain Recognition Task. *Proceedings of the 4th International Conference on Computer Recognition Systems*. Berlin, Heidelberg, 2005 :329–36.

19. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 1998; 2(2):121–67.

20. Pal M. Random forest classifier for remote sensing classification. *Int J Remote Sens* 2005; 26(1):217–22.

21. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. San Francisco, California., 2016 :785–94.

22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Proceedings of NIPS 2017*. Long Beach, CA, 2017 :4765–74.

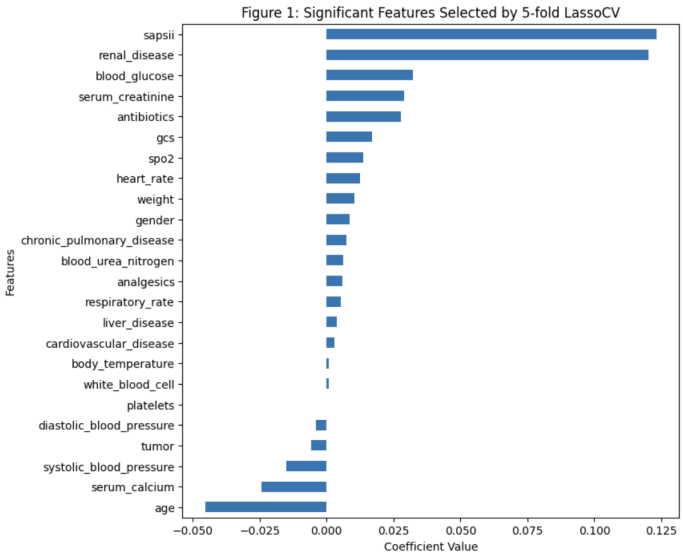
23. Basile JN, Bloch MJ. Overview of hypertension in adults [Internet]. Waltham (MA): UpToDate; 2024 [updated 2024 Oct 15; cited 2024 Oct 29]. Available from: `https://www.uptodate.com/contents/overview-of-hypertension-in-adults`

24. Mohamadlou H, Lynn-Palevsky A, Barton C, et al. Prediction of Acute Kidney Injury With a Machine Learning Algorithm Using Electronic Health Record Data. *Can J Kidney Health Dis* 2018; 5:2054358118776326

25. Koyner JL, Carey KA, Edelson DP, et al. The development of a machine learning inpatient acute kidney injury prediction model. *Crit Care Med* 2018; 46(7):1070–7

26. Kalisnik JM, Bauer A, Vogt FA, et al. Artificial intelligence-based early detection of acute kidney injury after cardiac surgery. *Eur J Cardiothorac Surg* 2022; 62(5):ezac289.

Appendix



GitHub Repository

https://github.com/lachlan-mcalpine-unimelb/COMP90089$_{G}roup13.git$