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.

Early Prediction of Acute Kidney Injury in Pediatric Intensive Care Unit Patients Using an XGBoost-Based Machine Learning Model: A Retrospective Study

Amir Ghadipasha
Emergency management research center, health management research institute, Iran university of medical sciences, Tehran, Iran
Mohammad Mehrnia
Iran University of Medical Sciences
Shiva Hosseinian
Shivahoseinian1@gmail.com
Iran University of Medical Sciences https://orcid.org/0009-0005-5543-3684
Sara Vafadar
Iran University of Medical Sciences
Zahra Hosseinzade
Iran University of Medical Sciences
Morteza Hemmat
Iran University of Medical Sciences
Negin Mousaeinejad
Iran University of Medical Sciences

Research Article

Keywords: Acute kidney injury, Pediatric intensive care unit, Machine learning, XGBoost, Early prediction

Posted Date: May 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9644930/v1

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

Abstract

Background

Acute kidney injury (AKI) is a frequent and serious complication in pediatric intensive care units (PICUs), and early detection is limited by delayed creatinine-based criteria. Machine learning (ML) using electronic health record (EHR) data may improve early risk prediction.

Objective

To develop and evaluate an explainable Extreme Gradient Boosting (XGBoost) model for early AKI prediction in a general PICU population and compare its performance with the Renal Angina Index (RAI).

Methods

This retrospective study included 513 children aged 1–18 years admitted to two tertiary PICUs between 2011 and 2023. AKI was defined by KDIGO serum creatinine criteria. Routinely collected clinical, laboratory, ventilatory, and medication variables recorded prior to AKI onset were used. Model performance was assessed using accuracy, sensitivity, F1-score, AUROC, and net reclassification improvement (NRI).

Results

The ML model achieved an AUROC of 0.888 with accuracy 0.742, outperforming RAI (AUROC 0.470). NRI favored the ML model (0.504). Key predictors included serum creatinine measures, hemoglobin, and mean arterial pressure.

Conclusions

An explainable XGBoost model accurately predicted AKI in critically ill children and substantially outperformed RAI, supporting its use for early PICU risk stratification.

1. Introduction

Acute kidney injury (AKI) is a common and serious complication in critically ill pediatric patients, with reported incidences affecting up to approximately one-quarter of pediatric critical care patients (1, 2) and associated with increased morbidity, mortality, and long-term risk of chronic kidney disease (CKD)(3, 4). Early identification remains challenging because AKI is still primarily diagnosed using serum creatinine and urine output criteria, which typically indicate functional impairment and may occur only

after kidney injury has progressed (5, 6). Studies further show that a significant proportion of pediatric AKI develops within the first 72 hours of PICU admission, narrowing the window for preventive intervention (7). Moreover, AKI pathophysiology in children is heterogeneous, influenced by hemodynamic instability, nephrotoxic exposures, sepsis, and underlying organ dysfunction, complicating prediction using traditional clinical tools alone (5, 8). This complexity makes AKI an ideal target for big-data analytics beyond traditional bedside assessment (9).

Recent advances in machine learning (ML) have improved early AKI prediction by enabling the analysis of complex physiologic and biochemical patterns that traditional statistical methods cannot adequately capture (5, 10). ML-based approaches, particularly gradient boosting methods, have demonstrated strong discriminative performance in identifying patients at risk for AKI across both pediatric and adult critical care settings (5, 11). Within pediatric populations, machine learning models have been applied in general PICU cohorts and have demonstrated good to excellent discriminative performance for AKI prediction (1, 12).

Among these techniques, Extreme Gradient Boosting (XGBoost) has shown particularly strong performance in multiple AKI prediction studies, including pediatric critical care and perioperative settings (10, 13).

Despite these advancements, gaps remain: most studies focus on narrowly defined populations, rely on limited feature sets, or do not examine early prediction windows essential for timely intervention. There is a need for ML models specifically optimized for the general PICU population using routinely available EHR data. Therefore, this study aims to develop an XGBoost-based model for early prediction of AKI in PICU patients, addressing a critical unmet need for timely and accurate risk stratification in pediatric critical care.

2. Methods

2.1. Study Design

This retrospective predictive-modeling study used secondary data extracted from the electronic health records (EHR) of pediatric patients admitted to the Pediatric Intensive Care Units (PICUs) of Ali Asghar and Firouzabadi Teaching Hospitals, affiliated with Iran University of Medical Sciences. The study period spanned 2011–2023 (1390–1402). Data extraction, AKI labeling, and model development were conducted using machine-learning methods.

2.2. Study Population and Sampling Strategy

The study population included patients aged 1–18 years admitted to the PICUs of the two hospitals during the study period. Demographic, clinical, laboratory, ventilatory, and medication-related variables were extracted from the EHR system.

Because this was a retrospective machine-learning study, sample size was determined pragmatically based on data availability and established recommendations for predictive modeling, considering the number of candidate predictors. With approximately 30 candidate predictors and commonly recommended subject-to-predictor ratios of 15–20:1, a minimum sample size of 500 patients was targeted. Accordingly, data from 513 eligible patients were included.

A census-based approach was used whereby all eligible encounters meeting inclusion criteria and without exclusion criteria were included.

2.3. Eligibility Criteria

Inclusion criteria

- PICU admission between 2011 and 2023
- Age 1–18 years at admission
- Availability of valid clinical and laboratory data in the EHR

Exclusion criteria

- Age < 1 year or > 18 years
- Missing or invalid age documentation
- AKI occurring within the first 12 hours of PICU admission
- PICU length of stay < 24 hours
- Multiple admissions for the same patient (only one encounter retained)
- For patients with multiple encounters and identical maximum AKI stage, the encounter with the highest AKI stage or longest length of stay was retained

2.4. Data Collection and Variables

Extracted variables were grouped into four categories:

1. **Vital signs**: systolic and diastolic blood pressure, mean arterial pressure, heart rate, respiratory rate, and oxygen saturation.
2. **Laboratory variables**: serum creatinine, blood urea nitrogen (BUN), lactate, C-reactive protein (CRP), hemoglobin, platelet count, and other routinely measured indices.
3. **Medication exposure**: nephrotoxic agents (e.g., aminoglycosides, vancomycin, cisplatin) and vasoactive medications.
4. **Ventilator parameters**: ventilation mode, respiratory rate, positive end-expiratory pressure (PEEP), tidal volume, and fraction of inspired oxygen (FiO_2).

AKI Definition and Index Time

AKI was defined according to Kidney Disease Improving Global Outcomes (KDIGO) serum creatinine criteria. Baseline creatinine was obtained from prior measurements when available; otherwise, age- and sex-adjusted normative values were used, consistent with established pediatric AKI research practices.

For AKI stages 2–3, AKI onset was defined as the time of the first creatinine value meeting KDIGO criteria. For AKI stage 1, onset was defined as the first creatinine value fulfilling stage 1 criteria. For non-AKI patients, a random reference time point during PICU admission was assigned.

All predictor variables were extracted strictly prior to the defined AKI onset or reference index time to prevent outcome leakage.

Renal Angina Index:

The Renal Angina Index (RAI) was calculated as a benchmark comparator but was intentionally excluded from model training to evaluate the standalone predictive performance of the machine learning model.

2.5. Model Development

An Extreme Gradient Boosting (XGBoost) algorithm was used to develop the AKI prediction model. XGBoost was selected for its ability to capture nonlinear interactions, handle missing values natively, and perform robustly on structured clinical data.

Data preprocessing:

Preprocessing included consistency checks, removal of duplicate records, and temporal alignment of time-stamped variables. No explicit imputation was performed; missing values were handled intrinsically by the XGBoost algorithm, which is designed to accommodate incomplete structured clinical data.

Feature selection:

Candidate predictors were selected based on clinical relevance and data availability. Final feature contributions were assessed using model-based importance measures during cross-validation.

Training and validation:

The dataset was split into training (80%) and test (20%) sets using stratified sampling to preserve the AKI outcome distribution. Hyperparameters were optimized using grid search with 5-fold cross-validation on the training set to prevent information leakage. The final model was trained using the optimized parameter set.

2.6. Statistical Analysis

Descriptive statistics were used to summarize patient characteristics. Categorical variables were reported as frequencies and percentages. Continuous variables were assessed for normality using

skewness and kurtosis. Normally distributed variables were summarized as mean ± standard deviation, and non-normally distributed variables as median with interquartile range (IQR).

Comparisons between patients with and without AKI were performed using independent t-tests or Mann–Whitney U tests for continuous variables and chi-square or Fisher's exact tests for categorical variables. These analyses were conducted for descriptive purposes only and were not used for feature selection or model training. Statistical significance was defined as $p < 0.05$. Analyses were performed using SPSS version 27.

2.7. Model Evaluation

Model performance was compared with the RAI using net reclassification improvement (NRI), applying a probability threshold of 0.5 for the machine-learning model and a cutoff value of 8 for RAI, in accordance with commonly used clinical thresholds.

Subgroup analyses were performed by sex and age group (infant, child, adolescent). Model performance was compared with the RAI using net reclassification improvement (NRI), with thresholds of 0.5 for the model and 8 for RAI.

Model explainability was assessed using SHapley Additive exPlanations (SHAP), partial dependence plots (PDP), and individual conditional expectation (ICE) plots. All machine learning analyses were conducted in Python (version ≥ 3.8) using pandas, scikit-learn, matplotlib, and seaborn. The final trained model was saved as a serialized file.

2.8. Ethical Considerations

The study protocol was approved by the Institutional Review Board and Ethics Committee of Iran University of Medical Sciences (IR.IUMS.FMD.REC.1403.225). Due to the retrospective design, informed consent was waived. All data were de-identified prior to analysis, and patient confidentiality was maintained in accordance with the Declaration of Helsinki.

3. Results

3.1. Study Population

A total of 513 pediatric patients admitted to the pediatric intensive care unit (PICU) were included in the analysis. The mean age was 4.73 ± 5.37 years, with a balanced sex distribution (51.9% female). Baseline demographic, hemodynamic, laboratory, and respiratory characteristics of the study population are summarized in Table 1. All predictor variables were obtained from measurements recorded before the first fulfillment of KDIGO criteria for AKI.

Exposure to nephrotoxic medications occurred in 35.9% of patients, 30.2% required mechanical ventilation, and 24.2% received vasoactive agents.

Table 1
Descriptive Statistics of Demographic and Clinical Variables in Pediatric Intensive Care Unit Patients (n = 513)

Variable	**Distribution (n = 513)**
Age, years (mean ± SD)	4.73 ± 5.37
Sex, n (%)	
Variable	**Distribution (n = 513)**
– Female	266 (51.85%)
– Male	247 (48.15%)
Birth weight, kg (mean ± SD)	3.24 ± 0.61
Systolic blood pressure, mmHg (mean ± SD)	100.94 ± 17.52
Diastolic blood pressure, mmHg (mean ± SD)	70.66 ± 17.31
Mean arterial pressure, mmHg (mean ± SD)	80.82 ± 17.08
Heart rate, bpm (mean ± SD)	107.47 ± 21.74
BUN, mg/dL (mean ± SD)	18.34 ± 7.00
Baseline creatinine, mg/dL (median, IQR)	1.60 (0.80–4.00)
Lactate, mmol/L (mean ± SD)	12.60 ± 3.82
CRP, mg/L (mean ± SD)	24.02 ± 16.19
Hemoglobin, g/dL (mean ± SD)	12.58 ± 1.41
Platelets, ×10³/μL (mean ± SD)	302.29 ± 87.72
Renal Angina Index (RAI), mean ± SD	11.76 ± 5.40
Cardiopulmonary bypass (CPB) duration, minutes (mean ± SD)	82.30 ± 20.71
Cross-clamp time, minutes (mean ± SD)	53.85 ± 15.83
Nephrotoxic medication, n (%)	
– No	329 (64.13%)
– Yes	184 (35.87%)
Mechanical ventilation, n (%)	
– No	358 (69.79%)
– Yes	155 (30.21%)

Variable	Distribution (n = 513)
Vasoactive medication, n (%)	
– No	389 (75.83%)
– Yes	124 (24.17%)
PIM3 (mean ± SD)	0.10 ± 0.06
PaO_2, mmHg (mean ± SD)	85.66 ± 14.62
FiO_2 (mean ± SD)	53.45 ± 26.59
Respiratory rate, breaths/min (mean ± SD)	29.91 ± 6.34
Tidal volume, mL/kg (mean ± SD)	7.78 ± 1.05
PEEP, cmH_2O (mean ± SD)	5.94 ± 1.49
Oxygen saturation, % (median, IQR)	97 (96–98)

3.2. Comparison Between Patients With and Without AKI

Baseline demographic characteristics, hemodynamic parameters, illness severity (PIM3), and routine laboratory indices did not differ significantly between patients with and without AKI (Table 2).

Serum lactate levels were modestly higher in patients without AKI (p = 0.026), with substantial overlap between groups.

Exposure to nephrotoxic medications, use of mechanical ventilation, vasoactive drug administration, and cardiopulmonary bypass characteristics were comparable between groups. Overall, no statistically significant differences were observed in baseline demographic, clinical, or laboratory variables between patients with and without AKI, except for lactate.

Table 2
Comparison of Demographic, Clinical, and Laboratory Variables Between Patients With and Without AKI

Variable	No AKI (Stage 0)	AKI (Stage 1–3)	P-value
Age, years (mean ± SD)	5.24 ± 4.73	5.51 ± 4.74	0.527
Sex, n (%)	0.134		
– Female	121 (44.98%)	126 (51.64%)	
– Male	148 (55.02%)	118 (48.36%)	
Birth weight, kg (mean ± SD)	3.62 ± 3.27	3.60 ± 3.20	0.220
Systolic BP, mmHg (mean ± SD)	101.07 ± 18.05	100.80 ± 17.00	0.863
Diastolic BP, mmHg (mean ± SD)	70.78 ± 18.06	70.54 ± 16.51	0.875
Mean arterial pressure, mmHg (mean ± SD)	81.01 ± 17.81	80.61 ± 16.30	0.790
Heart rate, bpm (mean ± SD)	107.83 ± 23.02	107.07 ± 20.31	0.689
BUN, mg/dL (mean ± SD)	18.72 ± 7.14	17.92 ± 6.86	0.193
Baseline creatinine, mg/dL (median, IQR)	1.83 (0.90–4.20)	1.30 (0.80–3.80)	0.101
RAI score (mean ± SD)	12.03 ± 5.30	11.48 ± 5.53	0.251
Lactate, mmol/L (mean ± SD)	12.96 ± 3.70	12.21 ± 3.92	**0.026**
CRP, mg/L (mean ± SD)	24.30 ± 16.32	23.72 ± 16.10	0.686
Hemoglobin, g/dL (mean ± SD)	12.68 ± 1.40	12.47 ± 1.42	0.083
Platelets, ×10³/μL (mean ± SD)	299.93 ± 83.94	304.88 ± 91.98	0.526
Nephrotoxic drug exposure, n (%)	0.409		
– No	168 (62.45%)	161 (65.98%)	
– Yes	101 (37.55%)	83 (34.02%)	
Mechanical ventilation, n (%)	0.847		
– No	189 (70.26%)	169 (69.26%)	
– Yes	80 (29.74%)	75 (30.74%)	
Variable	**No AKI (Stage 0)**	**AKI (Stage 1–3)**	**P-value**
Vasoactive drug use, n (%)	0.681		
– No	206 (76.58%)	183 (75.00%)	
– Yes	63 (23.42%)	61 (25.00%)	

Variable	No AKI (Stage 0)	AKI (Stage 1–3)	P-value
PIM3 (mean ± SD)	0.10 ± 0.06	0.10 ± 0.06	0.427
PaO_2, mmHg (mean ± SD)	85.57 ± 14.35	85.77 ± 14.98	0.880
FiO_2 (mean ± SD)	52.86 ± 26.07	54.10 ± 27.25	0.598
Respiratory rate, breaths/min (mean ± SD)	29.81 ± 6.41	30.01 ± 6.28	0.729
Tidal volume, mL/kg (mean ± SD)	7.76 ± 1.05	7.80 ± 1.05	0.664
PEEP, cmH_2O (mean ± SD)	5.99 ± 1.53	5.88 ± 1.47	0.656
Oxygen saturation, % (median, IQR)	97 (96–98)	97 (95.75–98)	0.681
CPB duration, min (mean ± SD)	81.42 ± 19.68	83.34 ± 22.31	0.670
Cross-clamp time, min (mean ± SD)	54.60 ± 17.07	52.98 ± 14.63	0.629

3.3. Comparison Between Low and High Renal Angina Index Groups

Patients were stratified according to Renal Angina Index (RAI) scores into low (< 8) and high (≥ 8) risk groups. Patients in the high RAI group had higher heart rate, blood urea nitrogen, baseline creatinine, hemoglobin, and platelet counts, while age, birth weight, and blood pressure parameters were similar between groups ($p < 0.05$ where applicable).

Nephrotoxic medication exposure was observed exclusively in the high RAI group (49.2% vs. 0%, $p < 0.001$). No significant differences were observed between groups with respect to mechanical ventilation or vasoactive medication use. Among respiratory variables, PaO_2 and tidal volume were higher, while positive end-expiratory pressure (PEEP) was lower in the high RAI group.

3.4. Performance of the Machine Learning Model

The machine learning model achieved an accuracy of 0.742, sensitivity of 0.733, and an F1 score of 0.733 in predicting AKI occurring within the predefined prediction window prior to KDIGO stage fulfillment. The area under the receiver operating characteristic curve (AUROC) was 0.888. The Brier score for model calibration was 0.146.

In comparison, the RAI score used as a standalone predictor showed lower performance, with an accuracy of 0.450, AUROC of 0.470, and Brier score of 0.307 (Table 3).

Table 3
Comparison of Performance Metrics Between the Machine Learning Model and the RAI Score for Predicting AKI

Metric	Machine Learning Model	RAI Only
Accuracy	0.742	0.450
Sensitivity (Recall)	0.733	0.689
F1 Score	0.733	0.544
AUROC	0.888	0.470
Brier Score	0.146	0.307

3.5. Feature Importance and Model Explainability

Feature importance analysis identified final serum creatinine measured prior to fulfillment of KDIGO criteria as the most influential predictor, followed by baseline creatinine, hemoglobin concentration, mean arterial pressure, and platelet count. Additional variables, including birth weight, heart rate, tidal volume, lactate level, and respiratory rate, contributed modestly to model performance.

Partial dependence and ICE analyses showed increasing predicted AKI risk with higher creatinine levels and a non-linear association for baseline creatinine. Hemoglobin, mean arterial pressure, and platelet count exhibited relatively stable effects across observed ranges.

SHapley Additive exPlanations (SHAP) analysis showed that creatinine-related variables had the strongest positive contributions to individual AKI risk, whereas hemoglobin and birth weight exhibited context-dependent effects through multivariable interactions rather than a uniform directional association.

3.6. Ablation Analysis and Clinical Utility

Ablation analysis demonstrated that removal of pre-AKI serum creatinine resulted in the greatest reduction in model performance (AUROC decrease of approximately 0.41). Removal of baseline creatinine and hemoglobin also led to notable declines in AUROC, whereas removal of several other variables resulted in minimal performance changes.

Decision curve analysis showed that the machine learning model provided greater net clinical benefit than treat-all or treat-none strategies across a wide range of threshold probabilities (0.05– 0.85). The model yielded a total net reclassification improvement (NRI) of 0.504 compared with the RAI score, with improvement driven primarily by correct reclassification of patients without AKI (NRI = 0.438).

4. Discussion

The present study aimed to develop and evaluate a machine learning model for early AKI prediction in children admitted to the PICU.

Our first main finding was that the developed ML model performed well in predicting AKI in critically ill children. Key performance metrics—including accuracy (0.742), sensitivity (0.733), F1-score (0.733), and AUROC (0.888)—indicate strong discriminative ability and good calibration in real-world clinical settings. These values were clearly superior to those of the Renal Angina Index (RAI), which is traditionally used as a risk stratification tool for AKI. The performance of our model is also comparable to that of other published ML-based AKI prediction models in PICU/ICU settings. For example, Hu et al. (2024) reported an explainable model in hospitalized children with an AUROC of 0.929, accuracy of 0.872, sensitivity of 0.886, and F1-score of 0.791, supporting the high effectiveness of EHR-based models in predicting AKI (12). Similarly, a systematic review by Du et al. (2023) showed that the sensitivity of most ML models for AKI prediction in ICUs ranges from 0.66 to 0.88, and accuracy typically from 0.70 to 0.87, depending on the algorithms used and the study population (5).

In adult ICUs, Zhang et al. (2025) reported an AUROC of 0.87 and F1-score of 0.76 using an explainable model, which is in line with our findings (10). Variations in performance across studies likely reflect differences in feature selection strategies, AKI labeling criteria, patient populations (adult vs pediatric), and the type and timing of variables included in each model.

The superior performance of the machine learning (ML) model compared with the Renal Angina Index (RAI) can be explained by several methodological and physiological factors. RAI is based on a small number of static baseline variables and cannot model nonlinear interactions or changes in clinical status over time. In contrast, the ML model incorporates repeated vital signs and laboratory measurements, allowing it to detect subtle patterns and evolving shifts that precede AKI onset. This ability to analyze dynamic information likely accounts for the markedly higher AUROC, accuracy, and sensitivity achieved by the ML model.

Feature importance analysis showed that final serum creatinine, baseline creatinine, and hemoglobin were the most influential predictors in our model. This pattern is consistent with prior pediatric ICU studies demonstrating the central role of both admission and longitudinal creatinine measurements in early AKI risk stratification (1, 6), as well as evidence supporting the contribution of hemoglobin to renal vulnerability in critically ill children (11).

Explainability analyses provided further insight into how the model uses clinical information. SHAP analysis showed that final creatinine had the largest positive impact on AKI risk, while baseline creatinine and hemoglobin also contributed meaningfully. This pattern was consistent with feature importance rankings and with the observed deterioration in AUROC when these key variables were removed, indicating good internal coherence of the model.

Similar findings have been reported in prior explainable ML studies in cardiac surgery (14), COPD (15), and general AKI prediction models (4).

Partial dependence and individual conditional expectation analyses revealed consistent patterns among secondary predictors. Higher BUN and lactate levels were associated with increased AKI risk, consistent

with prior ML-based studies (16, 17), while mean arterial pressure (MAP) demonstrated a U-shaped relationship in which both low and high values were associated with increased predicted risk (12).

4.1. Clinical Relevance

Decision curve analysis demonstrated that the ML model offers greater net clinical benefit than “treat all” or “treat none” strategies across a wide range of threshold probabilities, indicating that it may help guide practical decision-making, particularly for patients with intermediate risk. This is especially important in the PICU, where early recognition of high-risk patients can meaningfully influence outcomes.

Reclassification metrics further supported the clinical usefulness of the model. The overall net reclassification improvement (NRI = 0.504) showed that the model more accurately assigned patients to appropriate risk categories than the RAI, especially by correctly identifying children without AKI. Similar improvements in NRI have been reported in previous ML-based AKI studies using Random Forest, XGBoost, and LightGBM models, highlighting the advantage of data-driven approaches over traditional scoring systems.

Together, the DCA and NRI results indicate that the proposed model not only discriminates well but also improves real-world risk stratification, with the potential to reduce unnecessary interventions, optimize resource allocation, and support earlier preventive care in critically ill children.

In practice, such a model could be deployed as a background EHR risk monitor to identify intermediate-risk children who may benefit most from early preventive interventions.

4.2. Strengths and Limitations

This study has several notable strengths. It leveraged routinely collected EHR data and a modern, explainable ML framework, enabling analysis of dynamic clinical variables and providing transparent insight into feature contributions. The evaluation included multiple complementary metrics—discrimination, calibration, DCA, and NRI—offering a comprehensive assessment of model performance and potential clinical utility.

Several limitations should also be considered. First, external validation was not performed, and future studies are needed to assess model performance across diverse populations, institutions, and healthcare settings. Second, the model was developed using data from only two tertiary PICUs, which may limit generalizability to other clinical environments or patient groups. Third, the study relied on secondary EHR data, which may include recording errors, missing information, and limited temporal resolution for variables such as medication timing or rapid hemodynamic fluctuations. Fourth, kidney-specific biomarkers (e.g., NGAL, cystatin C, IL-18) were not available, and the model therefore relied exclusively on routinely collected clinical variables. Finally, although the model is explainable, ML outputs reflect associations rather than causal relationships and may be misinterpreted if used without appropriate clinical context.

5. Conclusion

This study aimed to develop and evaluate a machine learning model for the early prediction of acute kidney injury (AKI) in children admitted to the pediatric intensive care unit. The results demonstrated that the proposed model achieved strong performance across multiple dimensions—including accuracy, sensitivity, discrimination, calibration, and clinical utility— and significantly outperformed the traditional Renal Angina Index (RAI). Key predictors such as serum creatinine measures, hemoglobin, and mean arterial pressure contributed most substantially to the model's performance. Decision-analytic and classification assessments further confirmed the clinical value of the model in identifying high-risk patients. Overall, these findings indicate that machine learning algorithms can serve as effective tools for early risk stratification, enabling more timely therapeutic interventions and potentially improving clinical outcomes in pediatric critical care settings.

Declarations

Author Contributions

- Conceptualization and study design: A.G., N.M.
- Data extraction and curation: M.M., S.V., Z.H.
- Methodology and machine learning analysis: M.H., N.M.
- Statistical analysis: M.H.
- Manuscript drafting: S.H.
- Critical revision of the manuscript: All authors
- Final approval of the version to be published**:** All authors

Funding

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

Conflict of Interest

The authors declare no conflicts of interest.

Ethical Approval

The study protocol was approved by the Ethics Committee of Iran University of Medical Sciences

(Approval No. **IR.IUMS.FMD.REC.1403.225**).

Informed consent was waived due to the retrospective nature of the study.

Acknowledgements

The authors have nothing to report.

References

1. Dong J, Feng T, Thapa-Chhetry B, Cho BG, Shum T, Inwald DP et al (2021) Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care. Crit Care 25(1):288
2. Raina R, Nada A, Shah R, Aly H, Kadatane S, Abitbol C et al (2024) Artificial intelligence in early detection and prediction of pediatric/neonatal acute kidney injury: current status and future directions. Pediatr Nephrol 39(8):2309–2324
3. Ortiz-Soriano V, Kabir S, Claure-Del Granado R, Stromberg A, Toto RD, Moe OW et al (2022) Assessment of a modified renal angina index for AKI prediction in critically ill adults. Nephrol Dial Transpl 37(5):895–903
4. Kamel Rahimi A, van der Ghadimi M, Canfell OJ, Pole JD, Sullivan C et al (2023) Machine learning clinical prediction models for acute kidney injury: the impact of baseline creatinine on prediction efficacy. BMC Med Inf Decis Mak 23(1):207
5. Du YH, Guan CJ, Li LY, Gan P (2023) 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
6. Dauvin A, Donado C, Bachtiger P, Huang KC, Sauer CM, Ramazzotti D et al (2019) Machine learning can accurately predict pre-admission baseline hemoglobin and creatinine in intensive care patients. NPJ Digit Med 2:116
7. Sanchez-Pinto LN, Khemani RG (2016) Development of a Prediction Model of Early Acute Kidney Injury in Critically Ill Children Using Electronic Health Record Data. Pediatr Crit Care Med 17(6):508–515
8. Schrier RW, Wang W (2004) Acute renal failure and sepsis. N Engl J Med 351(2):15969
9. Bagshaw SM, Goldstein SL, Ronco C, Kellum JA (2016) Acute kidney injury in the era of big data: the 15(th) Consensus Conference of the Acute Dialysis Quality Initiative (ADQI). Can J Kidney Health Dis. ;3:5
10. Zhang L, Li M, Wang C, Zhang C, Wu H (2025) Prediction of acute kidney injury in intensive care unit patients based on interpretable machine learning. Digit Health 11:20552076241311173
11. Fragasso T, Raggi V, Passaro D, Tardella L, Lasinio GJ, Ricci Z (2023) Predicting acute kidney injury with an artificial intelligence-driven model in a pediatric cardiac intensive care unit. J Anesth Analg Crit Care 3(1):37
12. Hu J, Xu J, Li M, Jiang Z, Mao J, Feng L Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study. EClinicalMedicine. ;68:102409.
13. Luo XQ, Kang YX, Duan SB, Yan P, Song GB, Zhang NY et al (2024) Machine LearningBased Prediction of Acute Kidney Injury Following Pediatric Cardiac Surgery: Model Development and

Validation Study. J Med Internet Res. 2023;25:e41142

14. Jiang J, Liu X, Cheng Z, Liu Q, Xing W (2023) Interpretable machine learning models for early prediction of acute kidney injury after cardiac surgery. BMC Nephrol 24(1):326
15. Jiang S, Xu L, Wang X, Li C, Guan C, Che L et al (2025) Risk prediction for acute kidney disease and adverse outcomes in patients with chronic obstructive pulmonary disease: an interpretable machine learning approach. Ren Fail 47(1):2485475
16. Zhang R, Yin M, Jiang A, Zhang S, Xu X, Liu L (2024) Automated machine learning for early prediction of acute kidney injury in acute pancreatitis. BMC Med Inf Decis Mak 24(1):16
17. Peng X, Li L, Wang X, Zhang H A Machine Learning-Based Prediction Model for Acute Kidney Injury in (2022) Patients With Congestive Heart Failure. Front Cardiovasc Med 9:842873