



Machine learning approaches toward an understanding of acute kidney injury: current trends and future directions

Inyong Jeong¹, Nam-Jun Cho², Se-Jin Ahn¹, Hwamin Lee¹, and Hyo-Wook Gil²

¹Department of Medical Informatics, College of Medicine, Korea University, Seoul; ²Department of Internal Medicine, Soonchunhyang University Cheonan Hospital, Cheonan, Korea

Acute kidney injury (AKI) is a significant health challenge associated with adverse patient outcomes and substantial economic burdens. Many authors have sought to prevent and predict AKI. Here, we comprehensively review recent advances in the use of artificial intelligence (AI) to predict AKI, and the associated challenges. Although AI may detect AKI early and predict prognosis, integration of AI-based systems into clinical practice remains challenging. It is difficult to identify AKI patients using retrospective data; information preprocessing and the limitations of existing models pose problems. It is essential to embrace standardized labeling criteria and to form international multi-institutional collaborations that foster high-quality data collection. Additionally, existing constraints on the deployment of evolving AI technologies in real-world healthcare settings and enhancement of the reliabilities of AI outputs are crucial. Such efforts will improve the clinical applicability, performance, and reliability of AKI Clinical Support Systems, ultimately enhancing patient prognoses.

Keywords: Artificial intelligence; Acute kidney injury; Machine learning; Clinical decision support systems

INTRODUCTION

Acute kidney injury (AKI) poses significant health and socioeconomic issues, including extended hospital stays, high medical costs, and elevated mortality rates [1,2]. The current AKI definition using the serum creatinine level and urine output has evolved through various clinical trials [3-6]. The NGAL, KIM-1, and L-FABP levels may aid early detection and treatment [7-11]. However, the assays are costly and the utilities of these early marker levels unclear. Also, the levels do not aid the establishment of treatment plans after diagnosis; their clinical utilities are thus limited. Sequential tests are often not scheduled when AKI is not anticipated; the diagnostic contributions of these early markers are thus limited [11-13].

Recently, artificial intelligence (AI) has been used to tackle these limitations [14]. Several studies have used AI to detect AKI early, or to evaluate AKI progression and prognosis [15-18].

However, integration of AI-based AKI systems into clinical practice remains challenging. This review describes the current AI research trends in the AKI context, the challenges posed by model development, and the future integration of AI into AKI research and clinical applications.

THE CURRENT STATE OF AI IN TERMS OF AKI RESEARCH

We searched the Web of Science from 2014 to 2023 using the search terms “AKI” OR “Acute Kidney Injury” (Topic) and “Deep Learning” OR “Machine Learning” (Topic). The numbers of relevant publications exhibited a rapid increase in recent years (Fig. 1). The medical data are those of individual institutions and are often not shared because they are personal [19]. Consequently, we initially focused on information in large public databases including the electronic in-

tensive care unit (eICU) and the medical information mart for intensive care (MIMIC) [20-22]. Recent efforts have sought to refine predictions by the target group (patients in general wards or those undergoing specific surgeries); additional datasets have been built [23-26]. Most authors focused on

early AKI diagnosis but some attempted to predict prognosis [27-30]. The results were good; the areas under the receiver operating characteristic curves (AUROCs) ranged from 0.7 to 0.9 [31-33]. However, it is unclear whether such models are applicable in real clinical settings. We describe the trends in AI-related AKI research and the challenges in chapter 2.

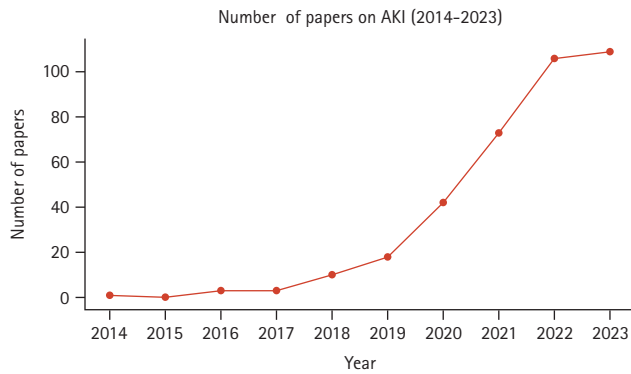


Figure 1. The annual numbers of acute kidney injury (AKI) artificial intelligence-related publications.

Labeling issues

Most researchers use the Kidney Disease: Improving Global Outcomes (KDIGO) criteria when diagnosing AKI; these practical guidelines are based on serum creatinine levels and urine output [34]. Consensus statements have been developed to define recovery of kidney function. Acute kidney disease is defined as AKI of stage 1 persisting for more than 7 days but less than 90 days, and chronic kidney disease is diagnosed when kidney function does not recover after 90 days [35,36]. Figure 2 presents a summary of the various AKI diagnostic criteria. Although these are relatively clear, there is no consensus on how to determine whether AKI

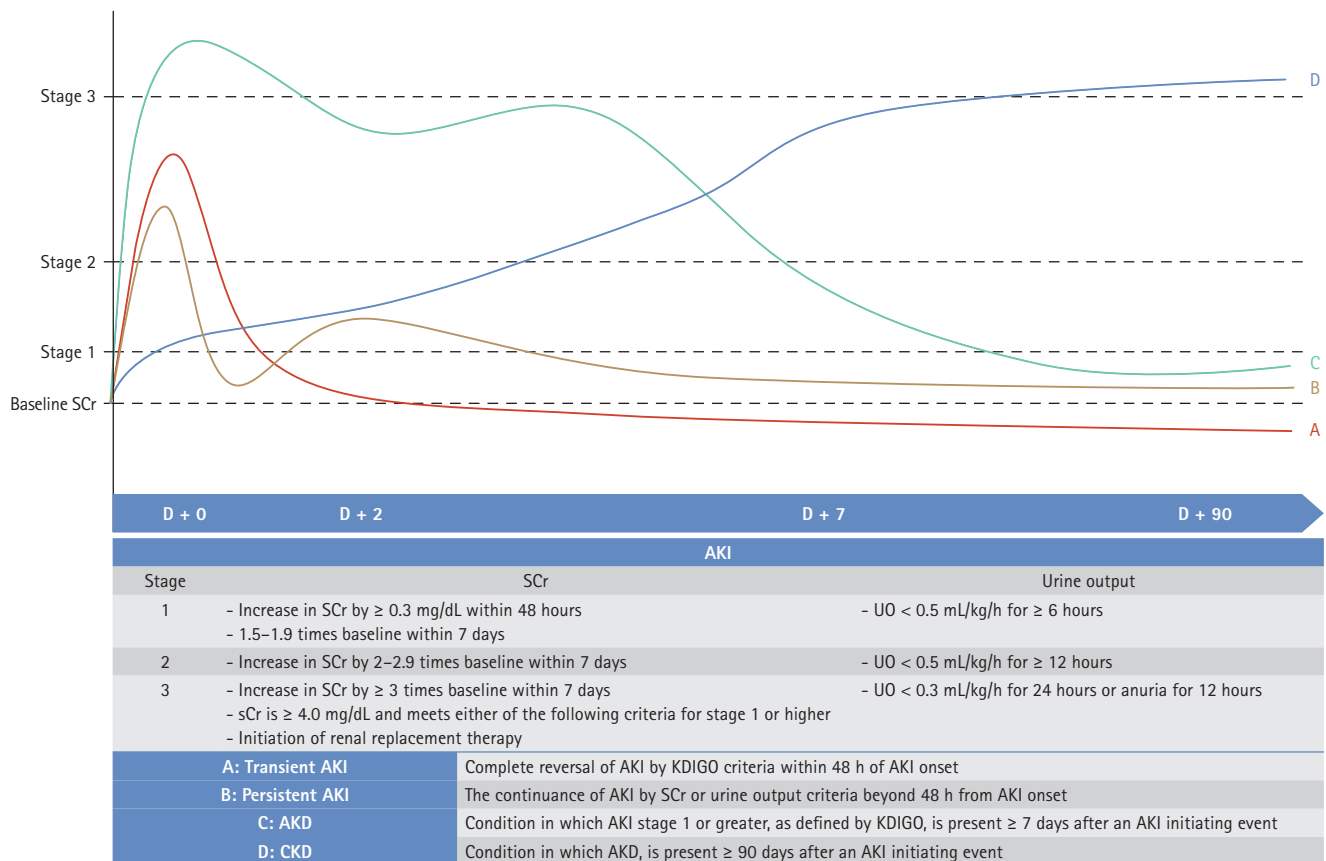


Figure 2. AKI definitions. AKI, acute kidney injury; SCr, serum creatinine; UO, urine output; AKD, acute kidney disease; CKD, chronic kidney disease; KDIGO, Kidney Disease: Improving Global Outcomes.

persists or if kidney function has recovered. Moreover, there are ongoing concerns regarding the use of serum creatinine levels and the estimated glomerular filtration rate (eGFR); it has been suggested that the GFR reserve (the stimulated GFR minus the basal GFR) might be more appropriate, but practical constraints arise when seeking to overcome the limitations of labeling [37-39].

Most machine learning models are developed and evaluated using retrospective data [40-42]. As such data do not track serum creatinine levels as systematically as do prospective studies, the data are incomplete. This limits our ability to specify the exact timing of AKI occurrence and to define recovery [43,44]. Furthermore, an operational definition based on serum creatinine levels and urine output renders patient identification rather ambiguous, raising various issues. When direct patient interaction is limited, the data are inadequate, and the various criteria are inconsistently used; model reliability is low. Thus, AI model development is challenging. Researchers have often used the most recent serum creatinine value before admission or specific surgery as the baseline level. If data are limited, the serum creatinine level on day 1 is used to predict AKI occurrence from day 2 onwards. When a model is to be applied in real-time, the baseline serum creatinine level is dynamically defined based on measurements taken within a specific time (e.g., 48 hours, 168 hours) or a measurement at a specific point, time is ignored (e.g., The most recent measurement before surgery, within the first 24 hours of admission). If multiple serum cre-

atinine measurements are available, the minimum, average, or most recent value is used [45-48]. Figure 3 shows the serum creatinine records of an actual patient. Depending on the baseline serum creatinine criterion and the research design, the patient may or may not have AKI. Accurate labeling is essential when training models. Although most research is based on the KDIGO criteria, the application of the criteria varies across studies, and this is associated with differences in the AKI status of the same patient; incorrect labeling distorts reality [49-51].

The Institutional Review Boards (IRBs) of Soonchunhyang University Cheonan Hospital approved our study protocol (approval numbers: 2019-10-023). The need for informed patient consent was waived by the IRBs as this was a retrospective review of anonymized clinical data.

Other criteria have been used when baseline serum creatinine data are not available. Sometimes, patients lacking such data are excluded. Alternatively, serum creatinine values measured over more than 7 days, or representative values obtained during the hospital stay, or back-calculations based on eGFRs are employed [52-55]. Urine output is rarely recorded in general wards, and only sometimes in intensive care units (ICUs); patients with AKI are thus not 'labeled' [56-59]. Of all evaluated AKI cases, only 11% met only the urine output criterion [60]. Moreover, even if the KDIGO criteria are met, it is sometimes difficult to diagnose AKI. Errors in serum creatinine measurements, variations among the measurements, or temporary changes in the levels render

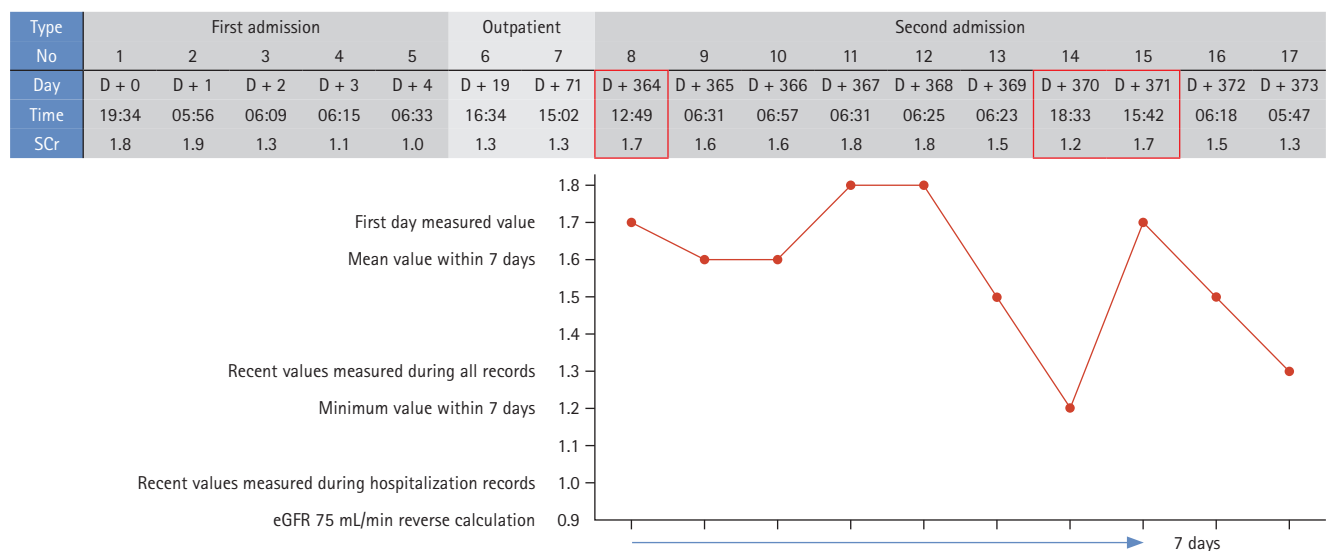


Figure 3. SCr levels and acute kidney injury diagnoses of real patients. SCr, serum creatinine; eGFR, estimated glomerular filtration rate.

it challenging to define the baseline serum creatinine level; the information is fragmented. However, to the best of our knowledge, such situations have not been considered.

Data preprocessing issues

It is essential that the population used for model development is similar to the population to which the model will be applied [61]. Data preprocessing, including the handling of missing and outlier information, is both essential and time-consuming [62,63]. Model development involves definition of the target group, anonymization, handling of outlier and missing data, labeling, and feature selection. The target group characteristics are significantly affected by preprocessing; it is important to describe how the model was developed. The methods used to locate and handle missing data, and the results, must be clearly presented. Such methods include deletion of features or subjects, use of the most recently measured value or a representative value, multiple imputation, model-specific methods, or combinations thereof [64-67]. The method must be practicable, reproducible, and not cause overfitting. For example, if a representative value during the hospital stay is used, how is that value chosen? It may be difficult to find an appropriate value, or the model may be overfitted by the training dataset [68,69]. Some previous studies did not appreciate the importance of missing data handling or did not clearly define the method used. Many works did not even clearly

describe the proportions of missing data before and after handling [27,31,70,71]. Care is essential when a model is to be applied in real-time. Depending on the method used to handle missing data, some patients may lack predictive points, leading to overestimations of patient numbers. If multiple imputation is employed, future information that should not be used at earlier points may be employed when handling missing data [72-74]. Figure 4 shows an example of how such situations can lead to overfitting and exaggerated patient numbers, attributable to data leakage during handling of time-series data.

Model evaluation issues

Machine learning models discover hidden patterns in training data but can become overfitted by the data [75,76]. Patient populations vary among hospitals, and the data characteristics vary by the equipment used and/or the prescribing practices of physicians [77,78]. Moreover, factors such as the frequencies of infectious diseases (example: COVID-19) change over time, and the continuous development of new drugs and surgical methods change the nature of the data over the years [79,80]. However, many studies employed only single cohorts and, to the best of our knowledge, very few models considered temporal changes [35,81,82].

Most model outputs are performance metrics. Performance is measured in various ways, including Accuracy, Precision, Recall, the F1 value, the AUROC, the area under

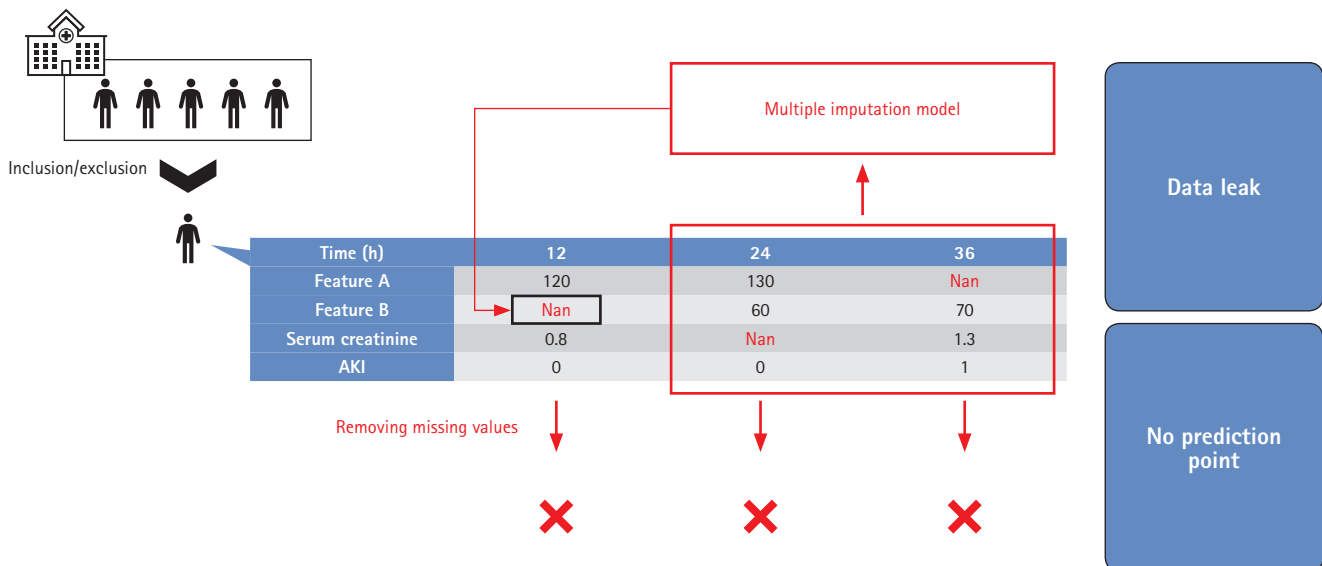


Figure 4. Overfitting and exaggeration of patient numbers caused by data leakage during processing of time-series data. AKI, acute kidney injury.

the precision-recall curve, among others, but the limitations mentioned above complicate the evaluation of models based solely on these metrics. The data used for performance evaluation, and the labeling methods and preprocessing methods, differ; it is meaningless to seek to compare model performance. Moreover, when handling a data imbalance during AKI model training, sampling not only of the training data but also the entire dataset can distort the performance metrics. For example, if data imbalance is severe, Recall and

Precision are important evaluation metrics. In such cases, as shown in Figure 5, the Precision varies by the sampling method. Table 1 summarizes the key points made above. Full model or code details were also considered important, but very few studies publish them [18,56,60,83,84]; these elements have been excluded from the Table. More references can be found in Supplementary Table 1.

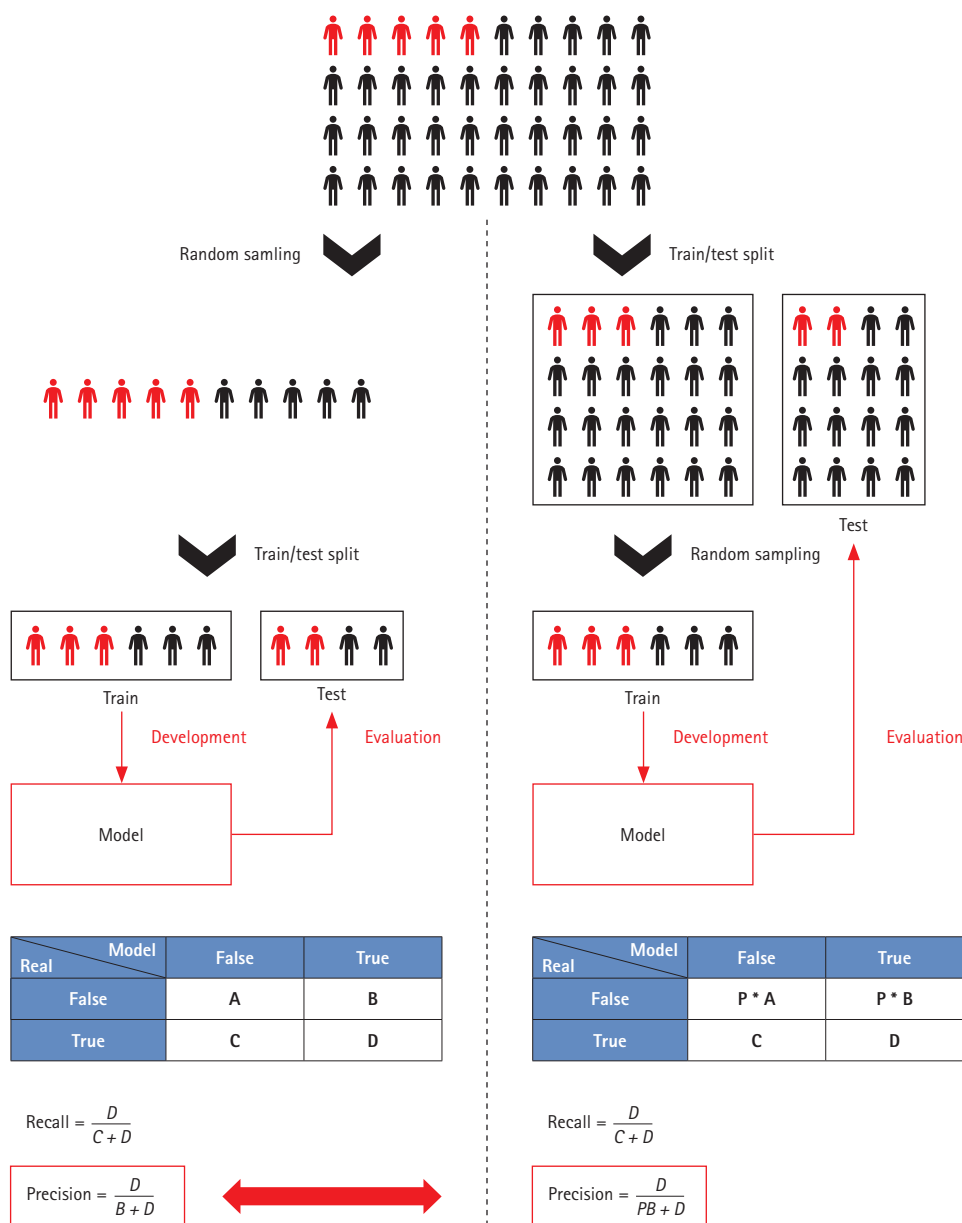


Figure 5. Variations in performance evaluation by the sampling method employed. Red, diseased populations; Black, non-diseased populations.

Table 1. Previous studies using artificial intelligence to predict acute kidney injury

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[15]	2018	AKI within 48 hours and RRT within 72 hours	The first value measured at the time of admission KDIGO without UO	After replacing with the most recent values within 12 hours, the remaining were replaced with median and mode	GBM	0.93	0	X	O	X
[46]	2019	Maximum SCr and AKI within 3 days in ICU	Measured value before ICU admission or minimum value on the first day of admission KDIGO with UO	After excluding features more than 20%, MICE was used	MLP	0.80	0	O	X	X
[40]	2018	AKI within 72 hours in general ward and ICU	Minimum value within 7 days or median within 365 days KDIGO with UO	Replaced with the most recent values	XGB	0.87	X	O	X	X
[83]	2019	AKI within 48 hours	Median of previous values or estimated value KDIGO with UO	Missing indicator was used	RNN	0.92	X	X	X	X
[16]	2020	AKI within 48 hours in emergency room, general ward, and ICU	The first value measured at the time of admission KDIGO without UO	After replacing with the most recent values within 12 hours, the remaining were replaced with median and mode	GBM	0.86	X	X	O	O
[18]	2020	AKI within 48 hours	Last measured value before admission or first measured value after admission KDIGO without UO	Replaced with the most recent values	GBM	0.93	X	O	O	O
[33]	2020	AKI occurrence in cardiac surgery patients (duration unspecified)	Not clearly stated KDIGO with UO	Not clearly stated	BN	0.85	x	x	△	O
[60]	2020	AKI within 7 days after cardiac thoracic surgery	Most recent value within 5 days before surgery KDIGO with UO	After replacing with the most recent values, the remaining were filled with the same value defined by doctor	RNN	0.89	X	X	O	X

Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[31]	2021	AKI within 12 hours in ICU	Minimum reported value during hospitalization AKIN with UO	SCr with the most recent value within 4 hours and urine output with the average within 9 hours	CNN	0.89	0	0	0	X
[32]	2021	AKI within 7 days in hospitalized patients and prediction SCr values over 3 days	The minimum value within 180 days before admission or the first value measured on the day of admission KDIGO without UO	Replaced with the most recent value	RNN	0.88	X	X	0	0
[47]	2021	AKI within 3 days in ICU	Minimum value within 7 days before ICU admission KDIGO without UO	Replaced with the most recent values	CNN	0.84	X	0	0	X
[48]	2021	AKI within 48 hours in ICU	Most recent value before admission or 20th percentile of a population with similar characteristics KDIGO with UO	After replacing with the most recent values, the remaining were filled with the median	CNN	0.86	X	0	X	X
[51]	2021	AKI within 7 days after liver cancer resection surgery	Measured value before surgery KDIGO without UO	Not clearly stated	RF	0.92	X	X	X	X
[50]	2021	AKI in patients who received organ donation in ICU	Not clearly stated KDIGO with UO	Not clearly stated	RF	0.85	X	X	X	X
[17]	2021	Survival of AKI patients in ICU without RRT	Not clearly stated	KNN imputer was used	LSTM	0.73	X	0	X	X
[24]	2022	AKI in septic patients in ICU	Estimated value KDIGO with UO	After excluding features with more than 20% missing values, multiple imputation was used	XGB	0.82	X	0	X	X
[52]	2022	AKI in acute cerebrovascular disease patients in ICU	Minimum value after ICU admission KDIGO without UO	Filled with 0	XGB	0.88	X	0	X	0

Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[84]	2022	AKI after partial nephrectomy	Measured value before surgery RIFLE and AKIN without UO	Not clearly stated	RF	0.75	X	X	X	X
[58]	2023	AKI recovery at discharge in stage 3 AKI patients admitted to the ICU	Most recent value within 3 months before admission or estimated value KDIGO without UO	After excluding features with more than 10%, the remaining were filled with mean and mode	LASSO	0.71	X	X	X	X
[27]	2022	Recovery of AKI at discharge in all patients with AKI	Most recent value within 7 days before admission or average within 90 days before admission KDIGO without UO	median considering AKI stage or KNN imputer was used	XGB	0.81	X	X	X	X
[26]	2022	Occurrence of AKI after immune checkpoint inhibitor administration in cancer patients	Most recent value within 3 months before the first ICPI KDIGO without UO	After excluding features with more than 20% missing values, MissForest was used	SVM	0.82	X	O	O	O
[56]	2023	AKI after brain surgery in ICU	Minimum value within 6 months before admission or estimated value KDIGO without UO	After excluding features with more than 20% missing values, multiple imputation was used	Ensemble	0.85	X	X	X	X
[28]	2023	Death, RRT, or reduction of eGFR by more than 50% in patients with AKI within 3 days in ICU	Measured value on the day closest to admission within 365 days before admission or estimated value KDIGO without UO	Replaced with mean value	RF	0.78	O	X	X	O

Cr, creatinine; AKI, acute kidney injury; AUROC, area under the receiver operating characteristic curve; RRT, renal replacement therapy; KDIGO, Kidney Disease: Improving Global Outcomes; UO, urine output; GBM, gradient boosting machine; SCr, serum creatinine; ICU, intensive care unit; MICE, multiple imputation by chained equations; MLP, multilayer perceptron; XGB, extreme gradient boosting; RNN, recurrent neural network; BN, bayesian network; CNN, convolutional neural network; RF, random forest; KNN, K-nearest neighbors; LSTM, long short-term memory; RIFLE, risk, injury, failure, loss of kidney function, and end-stage kidney disease; AKIN, acute kidney injury network; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; ICPI, immune checkpoint inhibitor; eGFR, estimated glomerular filtration rate. "Estimated value" refers to the estimation of parameters, such as eGFR, using methods like modification of diet in renal disease when baseline is unavailable. "Δ" signifies insufficient information for assessment. "Ensemble" refers to the combination of individual models through methods like voting, each developed separately.

FUTURE DIRECTIONS FOR AKI AI RESEARCH

Consensus labeling standards and evaluation data

Clinical guidelines for AKI diagnosis aside, it is clear that consensus AKI labeling standards are needed for machine learning models that use retrospective data. A scientific approach and verification of clinical effectiveness are essential when the numbers of serum creatinine measurements over time vary. The use of different standards renders it difficult to compare and evaluate models. Consistent, rational labeling guidelines are essential. Ideally, a multinational, multi-institutional dataset, labeled using agreed standards, would be employed for performance evaluation of all models. Large medical databases such as eICU and MIMIC have accelerated the development of medical AI [85-88].

Multicenter studies and use of the latest technologies

High-quality data are essential. Sparse data do not well-train models and may introduce bias toward specific groups or cause data overfitting [89,90]. When developing medical AI, consideration of patients with diverse characteristics, treated in various institutions, enhances model training [91]. However, medical data are difficult to obtain and use because of legal issues and the need for quality assurance [78,92]. Thus, federated and transfer learning are being actively researched. During federated learning, models or weights (but not data) are shared among institutions for model training.

Transfer learning takes a pre-trained model and develops a new model based on that model. Both methods largely lack data security issues and effectively train models even when the data are limited [93-96]. This ensures that the cohorts used for model development are diverse. Meta-learning can overcome problems associated with among-cohort differences during model development [97], and quantifies such differences and their impacts [18,98].

Currently, most research is focused on early AKI prediction. Only a few studies have used machine learning to predict AKI prognosis [99]. This may require rather long observation periods and/or extensive serum creatinine records. No consensus recovery criteria are available; it is difficult to predict AKI prognosis. Efforts to overcome these limitations include increasing the data volume via transfer and/or federated learning, and the use of multitask learning models [100]. It is important to monitor advances in machine learning continuously and apply them in the AKI research field. Figure 6 shows the latest AI technologies.

Model evaluation methods employing real users

Medical AI is meaningful only when it helps physicians and patients. Both time-based performance and external validations should be considered. All models should be thoroughly reviewed in terms of real-world effectiveness when predicting primary outcomes such as AKI occurrence or recovery, death, the hospital stay, and a need for ICU transfer. Comparing models to physicians who seek to solve the

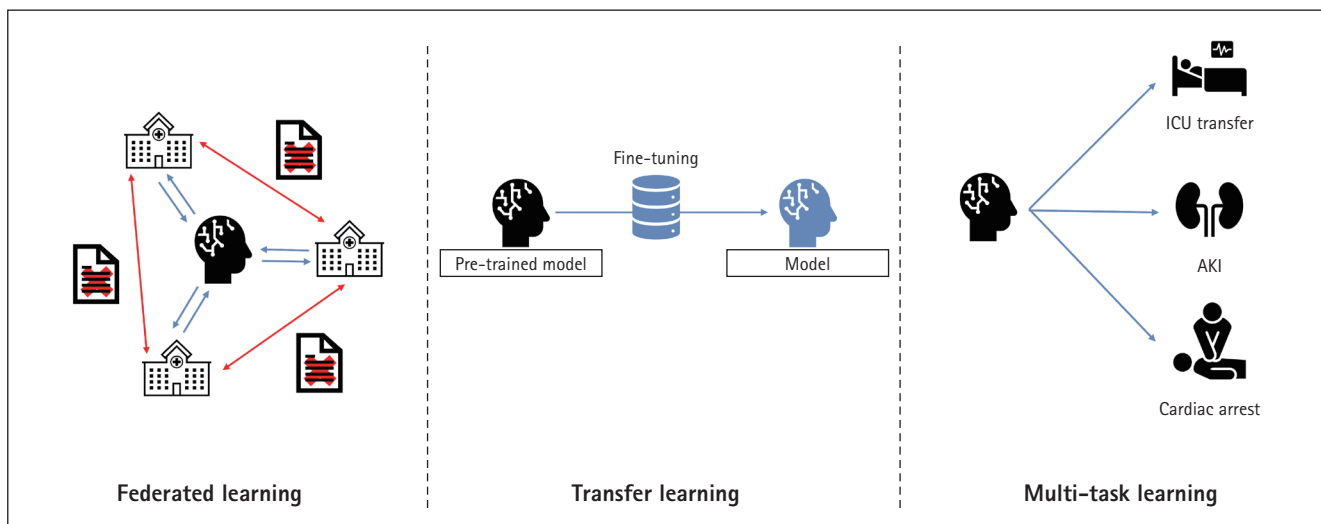


Figure 6. The latest artificial intelligence approaches. ICU, intensive care unit; AKI, acute kidney injury.

same tasks could be a useful alternative; this is frequently employed in the medical imaging AI field [101,102]. Rank et al. [60] found that AKI models outperformed physicians in terms of predicting AKI onset. However, tabular data are less informative than the information used by physicians [60]. Therefore, rather than comparing models with physicians, comparisons of physicians who do and do not use models might be better. Research on AI-based clinical decision support systems is ongoing.

Henry et al. [61] developed an AI model predicting sepsis onset, invited clinicians to use it, and conducted semi-structured interviews with 20 clinicians 6 months later. Prospec-

tive model evaluation after model introduction may be effective. Models should be continuously improved via user feedback and the results of prospective evaluations. The use of many features renders applications difficult; too much data are required [89,103]. However, if too few features are employed, a physician might have few options even if the model is accurate. Previous studies employed 15 to 1,000 features [18,104]. Useful features should be continuously sought based on model performance and user feedback. Figure 7 shows advanced model evaluation methods.

Explainable AI and causal inferences

Even if model performance is excellent, AI complexity (the 'black box problem') is a major obstacle to clinical applications [105]. The inner workings of a model are opaque; it is impossible to trace decision-making. In a 'white box' model, all operations are completely transparent, thus all steps from inputs to outputs. Many studies have sought to make black boxes white [106]. The advances range from an emphasis on simple feature importance to Local Interpretable Model-agnostic Explanations, SHapley Additive exPlanations, Partial Dependence Plots, Individual Conditional Expectation Plots, and combinations thereof [107]. However, these methods do not show how features cause certain de-

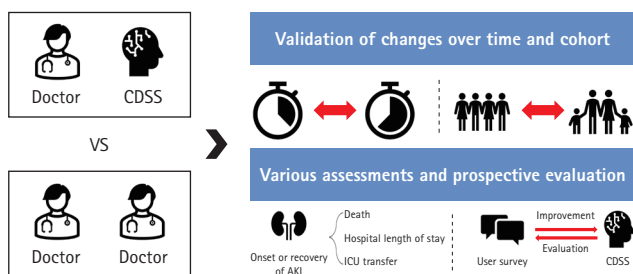


Figure 7. Several advanced models used for patient evaluation. CDSS, clinical decision support system; AKI, acute kidney injury; ICU, intensive care unit.

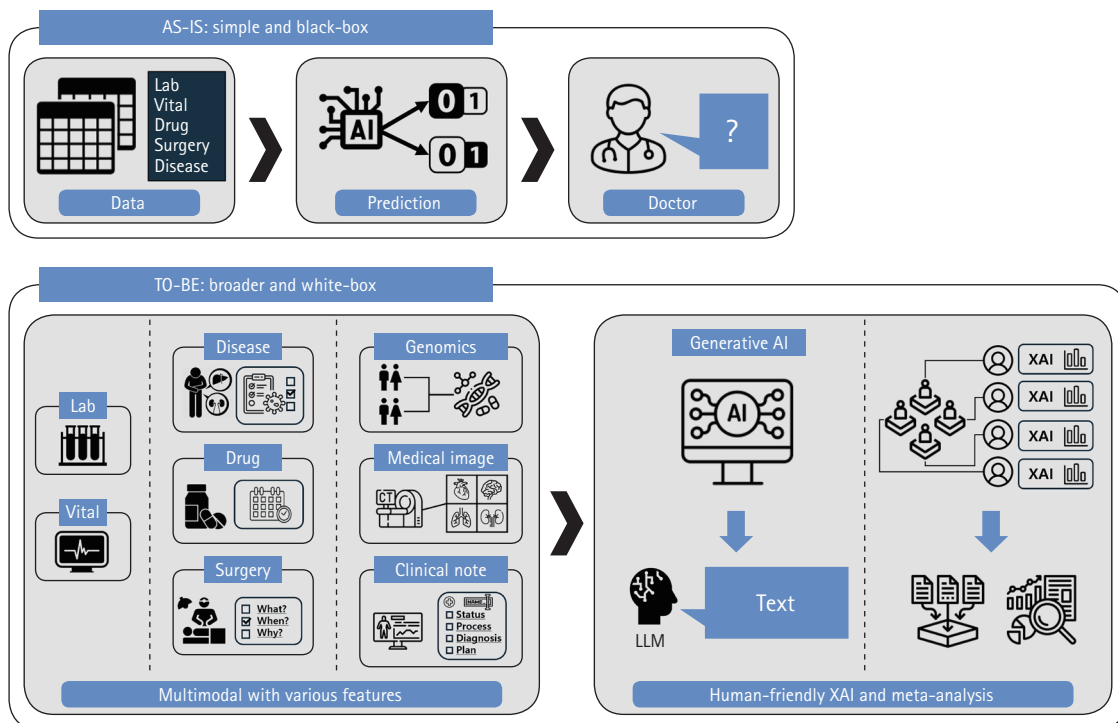


Figure 8. How to develop a reliable and clinically meaningful model. AI, artificial intelligence.

cisions to be made. Proof of causality, as opposed to proof of mere correlations or trends, is extremely challenging. It is crucial to derive correlations between various patient features and AKI [108,109]. Meta-analyses are aimed at strong associations between certain features and AKI occurrence or recovery, enhancing confidence in models that use such features [110]. AI researchers currently seek to create an Artificial General Intelligence that thinks like humans. A new AI should consider various modalities, not only fragmented information [111]. Although model interpretation may be difficult, multimodal technologies that combine tabular data with images or natural language, and generative AI technologies, could serve as alternatives to black boxes and prove difficult causalities [112]. Explanations of models using intuitive and human-friendly methods would greatly benefit users.

Simple laboratory parameters or vital signs can aid diagnoses and, thus, appropriate interventions by physicians [113]. For example, if AKI patients are identified early, physicians can adjust drug prescriptions or surgical schedules accordingly. This alleviates the burden on medical personnel who manage AKI and mitigates the differences between regions with good and limited medical infrastructures. Physicians must thoroughly review features that could enable more proactive interventions, such as diseases apparent on admission; the indications for the planned treatments/surgeries; and the medication types, quantities, and dosing schedules [114,115]. Consideration of such features during model development and interpretation would not only facilitate early diagnosis and intervention but also the development of an AI that truly supports clinical decision-making, thus beyond simple early detection (Fig. 8).

CONCLUSION

AI will greatly advance the practice of medicine. Many medical AI studies are in progress. In the AKI context, many works have reported high-level AI performance, confirming the potential of machine learning in terms of early AKI diagnosis and prediction of prognosis. However, medical AI is patient-intrusive; care is required. Many obstacles remain. It is time to move out of the computer laboratory into the clinic.

REFERENCES

1. Abebe A, Kumela K, Belay M, Kebede B, Wobie Y. Mortality and predictors of acute kidney injury in adults: a hospital-based prospective observational study. *Sci Rep* 2021;11:15672.
2. Rewa O, Bagshaw SM. Acute kidney injury-epidemiology, outcomes and economics. *Nat Rev Nephrol* 2014;10:193-207.
3. Jung HY, Lee JH, Park YJ, et al. Duration of anuria predicts recovery of renal function after acute kidney injury requiring continuous renal replacement therapy. *Korean J Intern Med* 2016;31:930-937.
4. Kellum JA, Sileanu FE, Murugan R, Lucko N, Shaw AD, Clermont G. Classifying AKI by urine output versus serum creatinine level. *J Am Soc Nephrol* 2015;26:2231-2238.
5. Allen JC, Gardner DS, Skinner H, Harvey D, Sharman A, Devonald MAJ. Definition of hourly urine output influences reported incidence and staging of acute kidney injury. *BMC Nephrol* 2020;21:19.
6. Kang C, Han SH, Park JS, Choi DE. Risk factors for post-contrast acute kidney injury in patients sequentially administered iodine- and gadolinium-based contrast media on the same visit to the emergency department: a retrospective study. *Kidney Res Clin Pract* 2023;42:358-369.
7. Lee K, Jang HR. Role of T cells in ischemic acute kidney injury and repair. *Korean J Intern Med* 2022;37:534-550.
8. Patschan D, Erfurt S, Oess S, et al. Biomarker-based prediction of survival and recovery of kidney function in acute kidney injury. *Kidney Blood Press Res* 2023;48:124-134.
9. Jung HH. Albuminuria, estimated glomerular filtration rate, and traditional predictors for composite cardiovascular and kidney outcome: a population-based cohort study in Korea. *Kidney Res Clin Pract* 2022;41:567-579.
10. Kim Y, Kang E, Chae DW, et al. Insufficient early renal recovery and progression to subsequent chronic kidney disease in living kidney donors. *Korean J Intern Med* 2022;37:1021-1030.
11. Wen Y, Parikh CR. Current concepts and advances in biomarkers of acute kidney injury. *Crit Rev Clin Lab Sci* 2021;58:354-368.
12. Kellum JA, Bihorac A. Artificial intelligence to predict AKI: is it a breakthrough? *Nat Rev Nephrol* 2019;15:663-664.
13. Yoo JJ, Park MY, Kim SG. Acute kidney injury in patients with acute-on-chronic liver failure: clinical significance and management. *Kidney Res Clin Pract* 2023;42:286-297.

14. Gheisari M, Ebrahimzadeh F, Rahimi M, et al. Deep learning: applications, architectures, models, tools, and frameworks: a comprehensive survey. *CAAI Trans Intell Technol* 2023;8:581-606.
15. Koyner JL, Carey KA, Edelson DP, Churpek MM. The development of a machine learning inpatient acute kidney injury prediction model. *Crit Care Med* 2018;46:1070-1077.
16. Churpek MM, Carey KA, Edelson DP, et al. Internal and external validation of a machine learning risk score for acute kidney injury. *JAMA Netw Open* 2020;3:e2012892.
17. Pattharanitima P, Vaid A, Jaladanki SK, et al. Comparison of approaches for prediction of renal replacement therapy-free survival in patients with acute kidney injury. *Blood Purif* 2021;50:621-627.
18. Song X, Yu ASL, Kellum JA, et al. Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. *Nat Commun* 2020;11:5668.
19. Wirth FN, Meurers T, Johns M, Prasser F. Privacy-preserving data sharing infrastructures for medical research: systematization and comparison. *BMC Med Inform Decis Mak* 2021;21:242.
20. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
21. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV [Internet]. *PhysioNet*, 2021 [cited 2023 Dec 22]. Available from: <https://physionet.org/content/mimiciv/1.0/>.
22. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data* 2018;5:180178.
23. Wei S, Zhang Y, Dong H, et al. Machine learning-based prediction model of acute kidney injury in patients with acute respiratory distress syndrome. *BMC Pulm Med* 2023;23:370.
24. 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;20:215.
25. Ko S, Jo C, Chang CB, et al. A web-based machine-learning algorithm predicting postoperative acute kidney injury after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2022;30:545-554.
26. Yu X, Wu R, Ji Y, Huang M, Feng Z. Identifying patients at risk of acute kidney injury among patients receiving immune checkpoint inhibitors: a machine learning approach. *Diagnostics (Basel)* 2022;12:3157.
27. Liu CL, Tain YL, Lin YC, Hsu CN. Prediction and Clinically important factors of acute kidney injury non-recovery. *Front Med (Lausanne)* 2022;8:789874.
28. Neyra JA, Ortiz-Soriano V, Liu LJ, et al. Prediction of mortality and major adverse kidney events in critically ill patients with acute kidney injury. *Am J Kidney Dis* 2023;81:36-47.
29. Jiang X, Hu Y, Guo S, Du C, Cheng X. Prediction of persistent acute kidney injury in postoperative intensive care unit patients using integrated machine learning: a retrospective cohort study. *Sci Rep* 2022;12:17134.
30. Li X, Wu R, Zhao W, et al. Machine learning algorithm to predict mortality in critically ill patients with sepsis-associated acute kidney injury. *Sci Rep* 2023;13:5223.
31. Alfieri F, Ancona A, Tripepi G, et al. A deep-learning model to continuously predict severe acute kidney injury based on urine output changes in critically ill patients. *J Nephrol* 2021;34:1875-1886.
32. Kim K, Yang H, Yi J, et al. Real-time clinical decision support based on recurrent neural networks for in-hospital acute kidney injury: external validation and model interpretation. *J Med Internet Res* 2021;23:e24120.
33. Li Y, Xu J, Wang Y, et al. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* 2020;43:752-761.
34. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* 2012;120:c179-c184.
35. Chawla LS, Bellomo R, Bihorac A, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup. *Nat Rev Nephrol* 2017;13:241-257.
36. Kung CW, Chou YH. Acute kidney disease: an overview of the epidemiology, pathophysiology, and management. *Kidney Res Clin Pract* 2023;42:686-699.
37. Lameire NH, Levin A, Kellum JA, et al. Harmonizing acute and chronic kidney disease definition and classification: report of a Kidney Disease: Improving Global Outcomes (KDIGO) Consensus Conference. *Kidney Int* 2021;100:516-526.
38. Kim H. The new race-free equations for estimating glomerular filtration rate: should they be adopted for Asians? *Kidney Res Clin Pract* 2023;42:670-671.
39. Lee YJ, Park YS, Park SJ, Jhang WK. Estimating baseline creatinine values to define acute kidney injury in critically ill pediatric patients. *Kidney Res Clin Pract* 2022;41:322-331.
40. 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.

41. Zhou Y, Feng J, Mei S, et al. Machine learning models for predicting acute kidney injury in patients with sepsis-associated acute respiratory distress syndrome. *Shock* 2023;59:352-359.
42. Jiang J, Liu X, Cheng Z, Liu Q, Xing W. Interpretable machine learning models for early prediction of acute kidney injury after cardiac surgery. *BMC Nephrol* 2023;24:326.
43. Zhang H, Wang AY, Wu S, et al. Artificial intelligence for the prediction of acute kidney injury during the perioperative period: systematic review and Meta-analysis of diagnostic test accuracy. *BMC Nephrol* 2022;23:405.
44. Kamel Rahimi A, Ghadimi M, van der Vegt AH, et al. Machine learning clinical prediction models for acute kidney injury: the impact of baseline creatinine on prediction efficacy. *BMC Med Inform Decis Mak* 2023;23:207.
45. 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 Int Conf Bioinformatics Biomed)* 2018;2018:683-686.
46. Zimmerman LP, Reyfman PA, Smith ADR, et al. Early prediction of acute kidney injury following ICU admission using a multivariate panel of physiological measurements. *BMC Med Inform Decis Mak* 2019;19(Suppl 1):16.
47. Sato N, Uchino E, Kojima R, Hiragi S, Yanagita M, Okuno Y. Prediction and visualization of acute kidney injury in intensive care unit using one-dimensional convolutional neural networks based on routinely collected data. *Comput Methods Programs Biomed* 2021;206:106129.
48. Le S, Allen A, Calvert J, et al. Convolutional Neural network model for intensive care unit acute kidney injury prediction. *Kidney Int Rep* 2021;6:1289-1298.
49. Zheng L, Lin Y, Fang K, Wu J, Zheng M. Derivation and validation of a risk score to predict acute kidney injury in critically ill cirrhotic patients. *Hepatol Res* 2023;53:701-712.
50. He ZL, Zhou JB, Liu ZK, et al. Application of machine learning models for predicting acute kidney injury following donation after cardiac death liver transplantation. *Hepatobiliary Pancreat Dis Int* 2021;20:222-231.
51. Dong JF, Xue Q, Chen T, et al. Machine learning approach to predict acute kidney injury after liver surgery. *World J Clin Cases* 2021;9:11255-11264.
52. Zhang X, Chen S, Lai K, Chen Z, Wan J, Xu Y. Machine learning for the prediction of acute kidney injury in critical care patients with acute cerebrovascular disease. *Ren Fail* 2022;44:43-53.
53. Rice ML, Barreto EF, Rule AD, et al. Development and validation of a model to predict acute kidney injury following high-dose methotrexate in patients with lymphoma. *Pharmacotherapy* 2024;44:4-12.
54. Ma Z, Liu W, Deng F, et al. An early warning model to predict acute kidney injury in sepsis patients with prior hypertension. *Heliyon* 2024;10:e24227.
55. Zulu C, Mwaba C, Wa Somwe S. The renal angina index accurately predicts low risk of developing severe acute kidney injury among children admitted to a low-resource pediatric intensive care unit. *Ren Fail* 2023;45:2252095.
56. Wu M, Jiang X, Du K, Xu Y, Zhang W. Ensemble machine learning algorithm for predicting acute kidney injury in patients admitted to the neurointensive care unit following brain surgery. *Sci Rep* 2023;13:6705.
57. Chen Q, Zhang Y, Zhang M, Li Z, Liu J. Application of machine learning algorithms to predict acute kidney injury in elderly orthopedic postoperative patients. *Clin Interv Aging* 2022;17:317-330.
58. Huang CY, Güiza F, De Vlieger G, et al. Development and validation of clinical prediction models for acute kidney injury recovery at hospital discharge in critically ill adults. *J Clin Monit Comput* 2023;37:113-125.
59. Sun S, Annadi RR, Chaudhri I, et al. Short- and long-term recovery after moderate/severe AKI in patients with and without COVID-19. *Kidney360* 2021;3:242-257.
60. Rank N, Pfahringer B, Kempfert J, et al. Deep-learning-based real-time prediction of acute kidney injury outperforms human predictive performance. *NPJ Digit Med* 2020;3:139.
61. Henry KE, Kornfield R, Sridharan A, et al. Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *NPJ Digit Med* 2022;5:97.
62. Vinisha FA, Sujihelen L. Study on missing values and outlier detection in concurrence with data quality enhancement for efficient data processing. *Proceedings of the 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT); 2022 Jan 20-22; Tirunelveli: IEEE, 2022: 1600-1607.*
63. Hoogland J, van Barneveld M, Debray TPA, et al. Handling missing predictor values when validating and applying a prediction model to new patients. *Stat Med* 2020;39:3591-3607.
64. Shawwa K, Ghosh E, Lanius S, Schwager E, Eshelman L, Kashani KB. Predicting acute kidney injury in critically ill patients using comorbid conditions utilizing machine learning. *Clin Kidney J* 2020;14:1428-1435.

65. Yang J, Peng H, Luo Y, Zhu T, Xie L. Explainable ensemble machine learning model for prediction of 28-day mortality risk in patients with sepsis-associated acute kidney injury. *Front Med (Lausanne)* 2023;10:1165129.
66. Peng C, Yang F, Li L, et al. A machine learning approach for the prediction of severe acute kidney injury following traumatic brain injury. *Neurocrit Care* 2023;38:335-344.
67. Luo XQ, Yan P, Zhang NY, et al. Machine learning for early discrimination between transient and persistent acute kidney injury in critically ill patients with sepsis. *Sci Rep* 2021;11:20269.
68. Perez-Lebel A, Varoquaux G, Le Morvan M, Josse J, Poline JB. Benchmarking missing-values approaches for predictive models on health databases. *Gigascience* 2022;11:giac013.
69. Liu Y, Qin S, Yepes AJ, Shao W, Zhang Z, Salim FD. Integrated convolutional and recurrent neural networks for health risk prediction using patient journey data with many missing values. *Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2022 Dec 6-8; Las Vegas (NV): IEEE, 2022: 1658-1663.*
70. Zhao X, Lu Y, Li S, 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. *Ren Fail* 2022;44:1326-1337.
71. Yang Y, Xiao W, Liu X, Zhang Y, Jin X, Li X. Machine learning-assisted ensemble analysis for the prediction of acute pancreatitis with acute kidney injury. *Int J Gen Med* 2022;15:5061-5072.
72. Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. *Sci Rep* 2018;8:6085.
73. Nijman SWJ, Hoogland J, Groenhof TKJ, et al. Real-time imputation of missing predictor values in clinical practice. *Eur Heart J Digit Health* 2020;2:154-164.
74. Nijman SWJ, Groenhof TKJ, Hoogland J, et al. Real-time imputation of missing predictor values improved the application of prediction models in daily practice. *J Clin Epidemiol* 2021;134:22-34.
75. Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. In: Montesinos López OA, Montesinos López A, Crossa J, ed. *Multivariate statistical machine learning methods for genomic prediction*. Cham: Springer, 2022:109-139.
76. Ying X. An overview of overfitting and its solutions. *J Phys Conf Ser* 2019;1168:022022.
77. Minvielle E, Fourcade A, Ricketts T, Waelli M. Current developments in delivering customized care: a scoping review. *BMC Health Serv Res* 2021;21:575.
78. Varoquaux G, Cheplygina V. Machine learning for medical imaging: methodological failures and recommendations for the future. *NPJ Digit Med* 2022;5:48.
79. Nadim MK, Forni LG, Mehta RL, et al. COVID-19-associated acute kidney injury: consensus report of the 25th Acute Disease Quality Initiative (ADQI) Workgroup. *Nat Rev Nephrol* 2020;16:747-764.
80. Bravata DM, Myers LJ, Perkins AJ, et al. Heterogeneity in COVID-19 patient volume, characteristics and outcomes across US Department of Veterans Affairs facilities: an observational cohort study. *BMJ Open* 2021;11:e044646.
81. Gao X, Ninan J, Bohman JK, et al. Extracorporeal membrane oxygenation and acute kidney injury: a single-center retrospective cohort. *Sci Rep* 2023;13:15112.
82. Zamirpour S, Hubbard AE, Feng J, Butte AJ, Pirracchio R, Bishara A. Development of a machine learning model of postoperative acute kidney injury using non-invasive time-sensitive intraoperative predictors. *Bioengineering (Basel)* 2023;10:932.
83. Tomašev N, Glorot X, Rae JW, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019;572:116-119.
84. Lazebnik T, Bahouth Z, Bunimovich-Mendrazitsky S, Halachmi S. Predicting acute kidney injury following open partial nephrectomy treatment using SAT-pruned explainable machine learning model. *BMC Med Inform Decis Mak* 2022;22:133.
85. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 2018;5:180161.
86. Wagner P, Strodthoff N, Boussejot RD, et al. PTB-XL, a large publicly available electrocardiography dataset. *Sci Data* 2020;7:154.
87. Yang J, Shi R, Wei D, et al. MedMNIST v2 - a large-scale lightweight benchmark for 2D and 3D biomedical image classification. *Sci Data* 2023;10:41.
88. Johnson AEW, Pollard TJ, Berkowitz SJ, et al. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data* 2019;6:317.
89. Jain A, Patel H, Nagalapatti L, et al. Overview and importance of data quality for machine learning tasks. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2020 Jul 6-10; New York (NY): Association for Computing Machinery, 2020:*

- 3561-3562.
90. Alzubaidi L, Bai J, Al-Sabaawi A, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *J Big Data* 2023;10:46.
 91. Lee SW, Lee HC, Suh J, et al. Multi-center validation of machine learning model for preoperative prediction of postoperative mortality. *NPJ Digit Med* 2022;5:91.
 92. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *NPJ Digit Med* 2020;3:119.
 93. Li L, Fan Y, Tse M, Lin KY. A review of applications in federated learning. *Comput Ind Eng* 2020;149:106854.
 94. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated learning for healthcare informatics. *J Healthc Inform Res* 2021;5:1-19.
 95. Jeong I, Kim Y, Cho NJ, Gil HW, Lee H. A novel method for medical predictive models in small data using out-of-distribution data and transfer learning. *Mathematics* 2024;12:237.
 96. Alzubaidi L, Al-Amidie M, Al-Asadi A, et al. Novel transfer learning approach for medical imaging with limited labeled data. *Cancers (Basel)* 2021;13:1590.
 97. Golriz Khatami S, Salimi Y, Hofmann-Apitius M, et al. Comparison and aggregation of event sequences across ten cohorts to describe the consensus biomarker evolution in Alzheimer's disease. *Alzheimers Res Ther* 2022;14:55.
 98. Zhao S, Sinha A, He Y, Perreault A, Song J, Ermon S. Comparing distributions by measuring differences that affect decision making. *Proceedings of the International Conference on Learning Representations* 2022; 2022 Apr 25-29; ICLR, 2022.
 99. Pickkers P, Darmon M, Hoste E, et al. Acute kidney injury in the critically ill: an updated review on pathophysiology and management. *Intensive Care Med* 2021;47:835-850.
 100. Zhang Y, Yang Q. A survey on multi-task learning. *IEEE Trans Knowl Data Eng* 2022;34:5586-5609.
 101. Deperlioglu O, Kose U, Gupta D, Khanna A, Giampaolo F, Fortino G. Explainable framework for Glaucoma diagnosis by image processing and convolutional neural network synergy: analysis with doctor evaluation. *Future Gener Comput Syst* 2022;129:152-169.
 102. Song W, Li S, Liu J, Qin H, Zhang B, Zhang S, Hao A. Multi-task cascade convolution neural networks for automatic thyroid nodule detection and recognition. *IEEE J Biomed Health Inform* 2019;23:1215-1224.
 103. Xu P, Ji X, Li M, Lu W. Small data machine learning in materials science. *npj Comput Mater* 2023;9:42.
 104. Hu Y, Liu K, Ho K, et al. A simpler machine learning model for acute kidney injury risk stratification in hospitalized patients. *J Clin Med* 2022;11:5688.
 105. Chan B. Black-box assisted medical decisions: AI power vs. ethical physician care. *Med Health Care Philos* 2023;26:285-292.
 106. Rai A. Explainable AI: from black box to glass box. *J Acad Mark Sci* 2020;48:137-141.
 107. Arrieta AB, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion* 2020;58:82-115.
 108. Roy S, Salimi. Causal inference in data analysis with applications to fairness and explanations. In: Bertossi L, Xiao G, ed. *Reasoning web. causality, explanations and declarative knowledge*. Cham: Springer, 2023:105-131.
 109. Nogueira AR, Pugnana A, Ruggieri S, Pedreschi D, Gama J. Methods and tools for causal discovery and causal inference. *Wiley Interdiscip Rev Data Min Knowl Discov* 2022;12:e1449.
 110. Shooshtari MM, Salehi M, Khalili H, Asadollahi E, Jamal-iMoghadamshahkhalil S. Investigating risk factors of AKI in patients with sepsis hospitalized in the intensive care unit. *Jundishapur J Chronic Dis Care* 2023;12:e130952.
 111. Bogdanovic Z. Artificial intelligence in federal information processing systems. *Am J Comput Sci Inform Technol* 2021;9:99.
 112. Hu J, Zhang Q, Yin H. Augmenting greybox fuzzing with generative AI. *arXiv:2306.06782 [Preprint]*. 2023 [cited 2023 Dec 22]. Available from: <https://doi.org/10.48550/arXiv.2306.06782>.
 113. Bora M, Staboliou E, Alexakou Z, et al. #5020 heterogeneity of hospital acquired acute kidney injury and the importance of early versus late nephrology assessment: a retrospective study of one center. *Nephrol Dial Transplant* 2023;38(Supplement 1):gfa063c_5020.
 114. Miller DD. The medical AI insurgency: what physicians must know about data to practice with intelligent machines. *npj Digit Med* 2019;2:62.
 115. Juma S, Goldszmidt M. What physicians reason about during admission case review. *Adv Health Sci Educ Theory Pract* 2017;22:691-711.

Received : March 21, 2024

Revised : April 26, 2024

Accepted : June 7, 2024

Correspondence to

Hyo-Wook Gil, M.D., Ph.D.

Department of Internal Medicine, Soonchunhyang University Cheonan Hospital, 31 Suncheonhyang 6-gil, Dongnam-gu, Cheonan 31151, Korea

Tel: +82-41-570-3682, Fax: +82-41-574-5762

E-mail: hwgil@schmc.ac.kr

<https://orcid.org/0000-0003-2550-2739>

CRediT authorship contributions

Inyong Jeong: resources, investigation, data curation, writing - original draft; Nam-Jun Cho: resources, investigation, data curation, writing - original draft; Se-Jin Ahn: data curation, visualization; Hwamin Lee:

conceptualization, methodology, investigation, data curation, writing - review & editing, funding acquisition; Hyo-Wook Gil: conceptualization, methodology, resources, investigation, data curation, writing - original draft, writing - review & editing, funding acquisition

Conflicts of interest

The authors disclose no conflicts.

Funding

This study was supported by the Soonchunhyang University Research Fund.

Supplementary Table 1. Summary of artificial intelligence in AKI assessment papers

Ref	Year	Task	Baseline Cr		Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
			Cr	AKI definition							
[S1]	2015	AKI within 2 to 9 days after admission	Average values within 7 to 365 days prior to admission	AKI definition: KDIGO without UO	Median imputation was used	LR	0.83	0	X	X	X
[S2]	2018	AKI at various periods after admission	The last measured value within 2 days prior to admission or the first measurement during hospitalization	KDIGO without UO	Removed	RF	0.77	X	X	△	X
[S3]	2018	AKI within 2 days post-liver transplantation	The most recent value measured before the surgery	AKIN without UO	For less than 1% missing, the mean and median of continuous variables; for less than 5% missing, hot-deck imputation was used	GBM	0.90	0	X	X	X
[S4]	2019	AKI at various periods after admission	The last measured value within 2 days before admission or the first measurement after admission	KDIGO without UO	Missing indicator was used	Ensemble	0.74	X	X	△	X
[S5]	2020	AKI and RRT within 48 hours in hospitalized patients	The first value measured at the time of admission	KDIGO without UO	After replacing with the most recent value, imputation with the median or mode was used	GBM	0.86	X	X	0	0
[S6]	2020	AKI within 7 days after aortic arch surgery	The most recent value before surgery	KDIGO without UO	Removed	LGBM	0.83	0	X	X	X
[S7]	2020	AKI within 7 days after cardiac surgery	The last measured value before surgery	KDIGO without UO	Replaced by mean and mode values	Ensemble	0.84	0	X	X	X
[S8]	2020	AKI within 7 days after partial nephrectomy	Not clearly stated	KDIGO with UO	Not clearly stated	XGB	0.74	X	X	X	X

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[S9]	2020	AKI after ICU admission	Average value within 7 to 180 days before admission KDIGO with UO	Replaced with mean values	GBM	0.69	X	X	X	O
[S10]	2022	AKI in patients undergoing knee arthroplasty	The last measured value within 6 months before surgery KDIGO without UO	Model's own algorithm was used	GBM	0.78	O	X	X	O
[S11]	2021	Survival within 1, 5, 17 years after kidney transplantation	Not clearly stated	Removed	SVM	0.82	X	X	X	X
[S12]	2021	AKI within 72 hours after ICU admission	The minimum value on the first day of admission KDIGO with UO	After removing rows with more than 50%, MICE was used	LGBM	0.90	X	O	X	X
[S13]	2021	AKI within 48 hours in PICU	Average considering age and gender KDIGO without UO	No missing values	Ensemble	0.89	X	X	O	O
[S14]	2021	AKI associated with cardiac surgery (duration not specified)	Not clearly stated Post-surgery SCr exceeding 2.26 mg/dL, doubling or more from the baseline, or RRT	After excluding features with more than 5%, patients with any missing values were removed	GBM	0.85	O	X	X	X
[S15]	2021	AKI within 7 days after nephrectomy in patients with renal cell carcinoma	Measured value before surgery (not clearly stated) KDIGO without UO	Not clearly stated	LGBM	0.81	X	X	X	X
[S16]	2021	Recovery of AKI within 48 hours in patients who developed AKI within 48 hours after ICU admission for sepsis	The minimum value during 7 days before admission KDIGO with UO	After excluding features with more than 30%, MissForest was used	ANN	0.76	O	O	X	X

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[S17]	2022	Severe AKI within 48 hours in the ICU	Not clearly stated KDIGO (AKI if stage II or III criteria are met, negative for the rest)	After removing features with more than 50%, 30-50% were checked with specialist and MICE was used for less than 30%	XGB	0.86	X	O	O	O
[S18]	2022	Prediction of progression from AKI stages 1/2 to AKI 3 in ICU patients	Not clearly stated KDIGO with UO	After removing features with more than 70%, XGBoost's internal method was used	XGB	0.93	X	O	X	X
[S19]	2022	AKI within 7 days after cardiac surgery	The first measured value at hospital admission KDIGO without UO	Some were filled in upon review, and multiple imputation was used for the remaining	DF	0.88	X	X	X	X
[S20]	2022	AKI within 7 days after cardiac surgery	The first measured value at hospital admission KDIGO without UO	Not clearly stated	RF	0.74	X	X	X	X
[S21]	2022	AKI within 7 days after liver transplantation in patients	The most recent value within 3 months before hospital admission, or the first measured value upon admission International Club of Ascites	Not clearly stated	ANN	0.81	X	X	X	X
[S22]	2022	Mortality during hospital stay in patients who developed AKI within 24 hours after ICU admission	The minimum value within 7 days KDIGO with UO	After removing features with more than 20%, multiple imputation was used	XGB	0.89	O	O	X	X
[S23]	2022	Recovery of AKI within 72 hours in patients who developed AKI within 48 hours after ICU admission	Not clearly stated	Not clearly stated	RF	0.77	O	X	X	X

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[S24]	2021	Oliguric AKI stages 2/3 in ICU	The minimum available value during the hospital stay or the most recent value within 4 days KDIGO with UO	Only urine volume and SCr were used, with SCr being the same as baseline and urine volume was calculated using a 9-hour average	ANN	0.91	0	X	0	0
[S25]	2022	Mortality during hospital stay in ICU patients over 60 years with AKD	The minimum value during the hospital stay or estimated value KDIGO with UO	Not clearly stated	XGB	0.90	0	X	X	0
[S26]	2022	AKI in patients hospitalized for acute pancreatitis	Not clearly stated KDIGO with UO	Not clearly stated	RF	0.91	X	X	X	△
[S27]	2022	AKI after orthopedic surgery in elderly patients	The most recent value before surgery KDIGO without UO	Not clearly stated	LR	0.82	X	X	X	X
[S28]	2022	Persistent AKI in patients who developed AKI after surgery in ICU	The minimum value within 6 months before ICU admission or estimated value KDIGO without UO	After removing features with more than 20%, multiple imputation was used	Emsemble	0.86	0	X	X	0
[S29]	2022	AKI in ICU for sepsis	Not clearly stated KDIGO with UO	After removing features with more than 20%, multiple imputation was used	XGB	0.82	0	X	X	X
[S30]	2021	Occurrence and recovery of AKI stages 2/3 in all patients	The minimum SCr during hospital stay KDIGO without UO (not clear)	Model's own algorithm was used	XGB	Not reported	X	0	X	X
[S31]	2023	Prediction of AKI stages 2/3 and AKI stage 3 after the initial 12 hours following Mechanical Ventilation	The average value before ICU admission or using the first measured value upon ICU admission KDIGO without UO (excluding criteria for acute dialysis, RRT)	Not clearly stated	RF	0.82	0	0	0	0

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[S32]	2024	Progression of AKI stages in patients who experienced AKI within 48 hours after cardiac surgery and received furosemide injection within 24 hours after AKI diagnosis	The minimum measured value within 3 months before cardiac surgery, excluding undefined patients KDIGO with UO	Multiple imputation was used	LR	0.81	0	X	X	X
[S33]	2023	Recovery of renal function within 1 year after utilizing RS in AKD patients	The minimum measured value within 365 days before RS, excluding undefined patients KDIGO without UO	Not clearly stated	RS	0.74	X	0	X	X
[S34]	2024	AKI within 7 days in lymphoma patients treated with methotrexate	The most recent value measured within 72 hours before medication administration, excluding undefined patients KDIGO without UO	Not clearly stated	LASSO	0.72	X	X	X	X
[S35]	2023	Prediction of stage 2 or higher AKI on the third day of hospitalization in pediatric ICU patients using RAI	The minimum value within 90 days before ICU admission or estimated value KDIGO with UO	Not clearly stated	RAI	0.77	X	X	X	X
[S36]	2023	AKI from 24 hours after admission until discharge in ICU-admitted patients with cirrhosis	Measured value within 7 days before ICU admission or the first measured value after ICU admission KDIGO with UO	After removing features with more than 80%, linear interpolation was used	ANN	0.76	0	X	X	0

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr AKI definition	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[S37]	2023	Mortality within 30 days, AKI recovery within 30 days, and AKI recovery within 90 days in AKI patients who received RRT	Not clearly stated KDIGO with UO	Not clearly stated	RF	0.67	X	X	X	X
[S38]	2023	Mortality within 7, 14, 28 days in patients who developed AKI within 24 hours after admission for sepsis to the ICU	Measured value before admission or the first value at admission KDIGO with UO	After removing features with more than 30%, multiple imputation was used	XGB	0.91	O	O	X	O
[S39]	2023	Mortality during hospital stay in patients who developed AKI within 48 hours after ICU admission for sepsis	Measured value before admission or the first value at admission KDIGO with UO	Multiple imputation was used	XGB	0.79	O	O	X	X
[S40]	2023	AKI in patients who developed ARDS on the first day after ICU admission	Applying ICD codes	After removing features with more than 20%, multiple imputation was used	XGB	0.87	O	X	X	X
[S41]	2023	Mortality within 28 days in patients who developed AKI on the first day after ICU admission for sepsis	Not clearly stated KDIGO with UO	After removing features with more than 20%, MissForest was used	XGB	0.87	X	O	X	X
[S42]	2023	Occurrence of stage 3 or higher AKI in TBI patients	Not clearly stated KDIGO with UO	Multiple imputation was used	RF	0.82	O	O	X	O

Supplementary Table 1. Continued

Ref	Year	Task	Baseline Cr	Methods for handling missing data	Model	AUROC	Presentation of missing data proportion	Using of public datasets	Real-time prediction	External validation
[543]	2024	Occurrence of stage 2 or higher AKI in septic patients with hypertension	Not clearly stated AKI definition	Not clearly stated	Nomogram	0.82	X	X	X	X

AKI, acute kidney injury; Cr, creatinine; AUROC, area under the receiver operating characteristic curve; KDIGO, Kidney Disease: Improving Global Outcomes; UO, urine output; LR, logistic regression; RF, random forest; AKIN, acute kidney injury network; GBM, gradient boosting machine; RRT, renal replacement therapy; LGBM, light gradient boosting machine; XGB, extreme gradient boosting; ICU, intensive care unit; SVM, support vector machine; PICU, pediatric intensive care unit; MICE, multiple imputation by chained equations; DF, deep forest; ANN, artificial neural network; SCr, serum creatinine; AKD, acute kidney disease; RS, renal scintigraphy; LASSO, least absolute shrinkage and selection operator; RAI, renal angina index; ARDS, acute respiratory distress syndrome; ICD, International Classification of Diseases; TBI, traumatic brain injury; eGFR, estimated glomerular filtration rate.

“Estimated value” refers to the estimation of parameters, such as eGFR, using methods like modification of diet in renal disease when baseline is unavailable. “Δ” signifies insufficient information for assessment. “Ensemble” refers to the combination of individual models through methods like voting, each developed separately.

SUPPLEMENTARY REFERENCES

- S1. Cronin RM, VanHouten JP, Siew ED, et al. National Veterans Health Administration inpatient risk stratification models for hospital-acquired acute kidney injury. *J Am Med Inform Assoc* 2015;22:1054-1071.
- S2. Cheng P, Waitman LR, Hu Y, Liu M. Predicting inpatient acute kidney injury over different time horizons: how early and accurate? *AMIA Annu Symp Proc* 2018;2017:565-574.
- S3. Lee HC, Yoon SB, Yang SM, et al. Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. *J Clin Med* 2018;7:428.
- S4. He J, Hu Y, Zhang X, Wu L, Waitman LR, Liu M. Multi-perspective predictive modeling for acute kidney injury in general hospital populations using electronic medical records. *JAMIA Open* 2019;2:115-122.
- S5. Churpek MM, Carey KA, Edelson DP, et al. Internal and external validation of a machine learning risk score for acute kidney injury. *JAMA Netw Open* 2020;3:e2012892.
- S6. Lei G, Wang G, Zhang C, Chen Y, Yang X. Using machine learning to predict acute kidney injury after aortic arch surgery. *J Cardiothorac Vasc Anesth* 2020;34:3321-3328.
- S7. Tseng PY, Chen YT, Wang CH, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care* 2020;24:478.
- S8. Zhu K, Song H, Zhang Z, et al. Acute kidney injury in solitary kidney patients after partial nephrectomy: incidence, risk factors and prediction. *Transl Androl Urol* 2020;9:1232-1243.
- S9. Shawwa K, Ghosh E, Lanius S, Schwager E, Eshelman L, Kashani KB. Predicting acute kidney injury in critically ill patients using comorbid conditions utilizing machine learning. *Clin Kidney J* 2020;14:1428-1435.
- S10. Ko S, Jo C, Chang CB, et al. A web-based machine-learning algorithm predicting postoperative acute kidney injury after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2022;30:545-554.
- S11. Naqvi SAA, Tennankore K, Vinson A, Roy PC, Abidi SSR. Predicting kidney graft survival using machine learning methods: prediction model development and feature significance analysis study. *J Med Internet Res* 2021;23:e26843.
- S12. Qian Q, Wu J, Wang J, Sun H, Yang L. Prediction models for AKI in ICU: a comparative study. *Int J Gen Med* 2021;14:623-632.
- S13. Dong J, Feng T, Thapa-Chhetry B, et al. Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care. *Crit Care* 2021;25:288.

- S14. Penny-Dimri JC, Bergmeir C, Reid CM, Williams-Spence J, Cochrane AD, Smith JA. Machine learning algorithms for predicting and risk profiling of cardiac surgery-associated acute kidney injury. *Semin Thorac Cardiovasc Surg* 2021;33:735-745.
- S15. Lee Y, Ryu J, Kang MW, et al. Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma. *Sci Rep* 2021;11:15704.
- S16. Luo XQ, Yan P, Zhang NY, et al. Machine learning for early discrimination between transient and persistent acute kidney injury in critically ill patients with sepsis. *Sci Rep* 2021;11:20269.
- S17. Liang Q, Xu Y, Zhou Y, et al. Severe acute kidney injury predicting model based on transcontinental databases: a single-centre prospective study. *BMJ Open* 2022;12:e054092.
- S18. Wei C, Zhang L, Feng Y, Ma A, Kang Y. Machine learning model for predicting acute kidney injury progression in critically ill patients. *BMC Med Inform Decis Mak* 2022;22:17.
- S19. Zhang H, Wang Z, Tang Y, et al. Prediction of acute kidney injury after cardiac surgery: model development using a Chinese electronic health record dataset. *J Transl Med* 2022;20:166.
- S20. Petrosyan Y, Mesana TG, Sun LY. Prediction of acute kidney injury risk after cardiac surgery: using a hybrid machine learning algorithm. *BMC Med Inform Decis Mak* 2022;22:137.
- S21. Bredt LC, Peres LAB, Risso M, Barros LCAL. Risk factors and prediction of acute kidney injury after liver transplantation: logistic regression and artificial neural network approaches. *World J Hepatol* 2022;14:570-582.
- S22. 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;20:2861-2870.
- S23. Zhao X, Lu Y, Li S, 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. *Ren Fail* 2022;44:1326-1337.
- S24. Alfieri F, Ancona A, Tripepi G, et al. A deep-learning model to continuously predict severe acute kidney injury based on urine output changes in critically ill patients. *J Nephrol* 2021;34:1875-1886.
- S25. Li M, Zhuang Q, Zhao S, et al. Development and deployment of interpretable machine-learning model for predicting in-hospital mortality in elderly patients with acute kidney disease. *Ren Fail* 2022;44:1886-1896.
- S26. Yang Y, Xiao W, Liu X, Zhang Y, Jin X, Li X. Machine learning-assisted ensemble analysis for the prediction of acute pancreatitis with acute kidney injury. *Int J Gen Med* 2022;15:5061-5072.
- S27. Chen Q, Zhang Y, Zhang M, Li Z, Liu J. Application of machine learning algorithms to predict acute kidney injury in elderly orthopedic postoperative patients. *Clin Interv Aging* 2022;17:317-330.
- S28. Jiang X, Hu Y, Guo S, Du C, Cheng X. Prediction of persistent acute kidney injury in postoperative intensive care unit patients using integrated machine learning: a retrospective cohort study. *Sci Rep* 2022;12:17134.
- S29. 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;20:215.
- S30. Sun S, Annadi RR, Chaudhri I, et al. Short- and long-term recovery after moderate/severe AKI in patients with and without COVID-19. *Kidney360* 2021;3:242-257.
- S31. Huang S, Teng Y, Du J, Zhou X, Duan F, Feng C. Internal and external validation of machine learning-assisted prediction models for mechanical ventilation-associated severe acute kidney injury. *Aust Crit Care* 2023;36:604-612.
- S32. Su Y, Zhang YJ, Tu GW, et al. Furosemide responsiveness predicts acute kidney injury progression after cardiac surgery. *Ann Thorac Surg* 2024;117:432-438.
- S33. Altarelli M, Jreige M, Prior JO, Nicod Lalonde M, Schneider AG. Renal scintigraphy to predict persistent renal failure after acute kidney injury: an observational study. *J Nephrol* 2023;36:1047-1058.
- S34. Rice ML, Barreto EF, Rule AD, et al. Development and validation of a model to predict acute kidney injury following high-dose methotrexate in patients with lymphoma. *Pharmacotherapy* 2024;44:4-12.
- S35. Zulu C, Mwaba C, Wa Somwe S. The renal angina index accurately predicts low risk of developing severe acute kidney injury among children admitted to a low-resource pediatric intensive care unit. *Ren Fail* 2023;45:2252095.
- S36. Zheng L, Lin Y, Fang K, Wu J, Zheng M. Derivation and validation of a risk score to predict acute kidney injury in critically ill cirrhotic patients. *Hepatol Res* 2023;53:701-712.
- S37. Wei W, Cai Z, Chen L, Yuan W, Fan Y, Rong S. Short-term prognostic models for severe acute kidney injury patients receiving prolonged intermittent renal replacement therapy based on machine learning. *BMC Med Inform Decis Mak* 2023;23:133.
- S38. Fan Z, Jiang J, Xiao C, et al. Construction and validation of prognostic models in critically ill patients with sepsis-associated acute kidney injury: interpretable machine learning

- approach. *J Transl Med* 2023;21:406.
- S39. Li X, Wu R, Zhao W, et al. Machine learning algorithm to predict mortality in critically ill patients with sepsis-associated acute kidney injury. *Sci Rep* 2023;13:5223.
- S40. Wei S, Zhang Y, Dong H, et al. Machine learning-based prediction model of acute kidney injury in patients with acute respiratory distress syndrome. *BMC Pulm Med* 2023;23:370.
- S41. Yang J, Peng H, Luo Y, Zhu T, Xie L. Explainable ensemble machine learning model for prediction of 28-day mortality risk in patients with sepsis-associated acute kidney injury. *Front Med (Lausanne)* 2023;10:1165129.
- S42. Peng C, Yang F, Li L, et al. A machine learning approach for the prediction of severe acute kidney injury following traumatic brain injury. *Neurocrit Care* 2023;38:335-344.
- S43. Ma Z, Liu W, Deng F, et al. An early warning model to predict acute kidney injury in sepsis patients with prior hypertension. *Heliyon* 2024;10:e24227.