BMC Medical Informatics and Decision Making

SYSTEMATIC REVIEW

Open Access

Exploring the role of Artificial Intelligence in Acute Kidney Injury management: a comprehensive review and future research agenda

Dima Tareq Al-Absi[1], Mecit Can Emre Simsekler[1*], Mohammed Atif Omar[1] and Siddiq Anwar[2,3]

Abstract

This study reviews the studies utilizing Artificial Intelligence (AI) and AI-driven tools and methods in managing Acute Kidney Injury (AKI). It categorizes the studies according to medical specialties, analyses the gaps in the existing research, and identifies opportunities for future research directions. PRISMA guidelines were adopted using the three most common databases (PubMed, Scopus, and EBSCO), which resulted in 27 eligible studies, published between 2012 and 2023. The study showed significant heterogeneity in the design of the models, with variations in clinical settings, patient characteristics, cohort regions, and statistical methods. Most models were developed for AKI in hospitalized patients, particularly those undergoing surgery or in intensive care units. Compact models with a subset of significant predictors were deemed more clinically applicable than full models with all predictors. The findings suggest that AI tools, such as machine learning (ML) algorithms, have high prediction capabilities despite the dynamic and complex association among the influencing factors and AKI. Based on these findings and the recognized need for broader inclusivity, future studies should consider adopting a more inclusive approach by incorporating diverse healthcare settings, including resource-limited or developing countries. This inclusivity will lead to a more holistic understanding of AKI management challenges and facilitate the development of adaptable and universally applicable AI-driven solutions. Additionally, further investigations should focus on refining AI models to enhance their accuracy and interpretability, promoting seamless integration and implementation of AI-based tools in real-world clinical practice. Addressing these key aspects will elevate the effectiveness and impact of AI-driven approaches in managing AKI.

Keywords Acute Kidney Injury, AKI, Patient safety, Artificial Intelligence, Healthcare operations, Healthcare analytics, Machine learning

*Correspondence:
Mecit Can Emre Simsekler
emre.simsekler@ku.ac.ae
[1]Department of Management Science and Engineering, Khalifa University of Science and Technology, P.O. Box 127788, Abu Dhabi, United Arab Emirates
[2]Department of Medicine, Sheikh Shakhbout Medical City, Abu Dhabi, United Arab Emirates
[3]College of Medicine and Health Science of Medicine, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates

Introduction

Acute kidney injury

Acute Kidney Injury (AKI) denotes a spectrum of syndromes characterized by a persistent decline in glomerular filtration rate (GFR) [1]. It is defined as a rapid deterioration in renal functions occurring for hours to days and leading to retention of blood urea Nitrogen waste (BUN) and elevation of serum creatinine (SCr) in the body [2]. AKI affects around 200 patients per million population, 22% of hospitalized patients, and more than 50% of patients in the intensive care unit (ICU) [3]. The development of AKI is associated with prolonged hospital stays and increased health-related expenditure [4]. A study by Chertow et al. revealed that the development of AKI leads to an increased hospital stay and higher hospitalization costs by an average of $7,500 [5].

Notably, AKI is associated with high mortality and morbidity rates; it is estimated that more than 2 million deaths have been reported indirectly to AKI every year [6]. Furthermore, patients who experience AKI during their hospital stay have nearly double the risk of death compared to those who do not develop AKI [7]. Besides the high mortality rate, economic costs, and resource usage associated with AKI, other severe consequences extend beyond hospital discharge [8, 9]. AKI survivors are at greater risk of progressing future health problems, such as chronic kidney disease (CKD), cardiovascular and cerebrovascular, and respiratory complications [10]. Additionally, AKI has been linked to acute kidney disease (AKD) and end-stage kidney disease, suggesting that even brief episodes of AKI can lead to significant long-term morbidity and increased risk of mortality [11, 12]. These health problems can significantly impact survivors' quality of life and often require long-term management, placing additional demands on healthcare resources.

In addition, recent studies have demonstrated that the transition of care from inpatient to outpatient requires special attention. When patients are discharged, they require prolonged care to recover and manage their kidney functions, which often includes staying in skilled nursing facilities or receiving outpatient follow-up care [11, 13]. This transition of care is important to stabilize their condition and prevent any further decline in kidney function. The prolonged recovery period impacts not only the patient, who may face significant lifestyle adjustments such as dietary restrictions, medication management, and reduced physical activity but also places a considerable demand on healthcare resources [14]. Skilled nursing facilities, regular follow-up visits, and specialized supportive services are needed to ensure proper monitoring, mitigate complications, and prevent readmissions. These ongoing demands increase the burden on healthcare systems, requiring additional staffing, infrastructure, and financial resources, which collectively contribute to the overall cost of AKI management [12, 15].

Further, AKI survivors have a higher risk of 30-day rehospitalization [16]. The minority (4%) of AKI survivors require referral to nephrology care within 60 days of discharge and (9%) within one year [17]. Similarly, according to the United States Renal Data System (USRDS), Annual Report 2015, (19%) of AKI survivors were referred for nephrology care one year after an AKI episode. These figures suggest that AKI incidence accounts for high personal and community costs and burdens the healthcare system [18, 19]. One way to reduce this growing burden is to devote efforts to AKI management.

AKI management

AKI management refers to the strategies and practices used to improve the detection, prevention, and management of AKI in clinical settings [20]. The management practices might also include the use of predictive models and algorithms. This approach has the potential to identify high-risk patients, implement protocols and interventions to manage AKI, utilize electronic health record (EHR) systems, a digital data format about patients' health, to monitor AKI-related data and educate medical providers on best practices for treating AKI. AKI operations management aims to improve patient outcomes and reduce the burden on the healthcare systems [21].

Timely recognition, early intervention, and appropriate follow-up are essential in managing AKI. Delayed recognition and inadequate management of AKI are associated with significantly poor patient outcomes [22]. These issues present challenges for healthcare systems and pose risks to patient safety. A study by Yang et al. revealed that the non-recognition of AKI in China was 74.2%, of which 17.6% were given a delayed diagnosis [23]. In an attempt to improve the diagnosis and treatment of AKI globally, The International Society of Nephrology (ISN) launched the "0by25" campaign with the target of eliminating avoidable AKI-related deaths [24]. A key part of achieving this goal is predicting AKI before it occurs and taking preventive measures.

Predicting AKI is important for managing prevention strategies effectively, especially for high-risk individuals [25]. Accurate and early risk assessment may help inform patients and their families about potential risks and guide appropriate management of AKI. While significant progress has been made in predicting and managing AKI, gaps still exist, highlighting the need for further research to determine clinical predictors that can efficiently identify those at high risk.

Recent studies have shown that AKI can be both predictable and preventable if early risk factors are identified [24]. Recognizing these factors is essential for effectively managing AKI to achieve better clinical outcomes.

Although AKI risk factors are complex and often overlap [26], accurately identifying them provides an opportunity to prevent occurrences and reduce negative outcomes associated with AKI by analyzing their correlation and impact through predictive models.

Risk prediction models in AKI

Risk prediction models (RPMs) are mathematical tools that use a combination of variables to predict the risk of an adverse event based on existing data [27]. RPMs can be applied in many fields, including healthcare.

RPMs can be both diagnostic and prognostic. Diagnostic RPMs estimate the likelihood of an existing disease, while prognostic RPMs assess the risk of a future disease for patients who currently do not have it. Powerful algorithms and advanced tools, such as Artificial Intelligence (AI) and its subset, machine learning (ML), are commonly used to develop accurate and reliable RPMs.

ML entails teaching a machine to learn from data patterns, and improve automatically, from experience, rather than being explicitly programmed with the possible scenarios as in traditional programming [28]. ML tools include decision trees, neural networks, support vector machines, and random forests among others [29]. ML solves prediction problems involving big datasets with enormous numbers of predictors [30]. Moreover, ML algorithms offer improved performance, robustness, and ease of clinical use [31].

Concerns have been raised regarding the limited predictive power of RPMs. One key issue in ML is its heavy reliance on the quality and relevance of input data for training. ML models often require large datasets to achieve acceptable performance levels. Moreover, deficiencies in the dataset can substantially lead to deficiencies in the model itself [32]. Therefore, the performance of a predictive model should be thoroughly assessed.

It is important to note that the potential impact of clinical RPMs on patient safety relies heavily on the quality and reliability of the model and the risk variables being utilized. Accurate models can effectively stratify patients to identify those at high risk of AKI development and thus provide specialized care to prevent additional renal insult. Clinical implementation of RPMs focuses on predicting AKI incidences early enough for relevant interventions to improve patient outcomes. Much research has been devoted to predicting AKI, where prevention strategies are considered for those at elevated risk. The advancements of ML have resulted in significant improvements in the RPMs used for estimating AKI incidences. Moreover, the increased availability of EHR made it possible to develop RPMs to estimate AKI risk factors [33].

The development of AKI RPMs is relevant to patients and their families, healthcare providers, policymakers, and medical researchers. Such models are essential to predict the outcome of AKI patients and are often used for risk stratification. Many models have been developed for AKI predictions and are widely used in diverse settings and populations [4, 8, 9]. Therefore, existing RPMs are based on different combinations of risk variables based on specific conditions.

Recent studies have demonstrated no standard combination of risk factors or a clear understanding of their impact and association with AKI episodes [34]. A study by Joseph et al. revealed that risk factors identified in some studies have not been confirmed, or their effects have not been consistent in subsequent studies [35]. Other studies revealed that existing RPMs are of poor utility due to the difference in the independent risk factors used to construct the models [36, 37].

However, medical providers claim that AKI tends to occur in those with common risk factors and certain conditions and have extensively researched these factors [34]. Moreover, introducing EHRs signifies a big step forward in predicting AKI risk enablers. A recent study demonstrated the critical role of EHRs in discovering the interactions between risk factors, reducing medical errors, and improving patient outcomes [38].

Given the impact of AKI prognosis, enhancing the early prediction of AKI incidence is essential to improve preventive, diagnostic, and intervention strategies. Attaining a comprehensive understanding of AKI risk prediction offers the potential to derive valuable insights that could maximize the clinical effectiveness of AKI prediction models. In contrast to previous SLR papers on AKI, this study goes beyond identifying common risk factors and conditions associated with AKI development. While the primary objective remains the identification of high-risk factors, we take a comprehensive approach by thoroughly reviewing existing models, exploring their characteristics and applications, examining the AI tools employed in detail, and comparing performance metrics across different models. By examining the specific attributes and functionalities of these RPMs, we aim to provide a better understanding of their strengths and limitations. Additionally, our study assesses the practicality and applicability of these models in real-world scenarios. This detailed analysis of the existing literature allows us to gain valuable insights that may contribute to the improvement of future prediction models for AKI. Furthermore, by highlighting the different ML tools used and their respective performances, we show the potential of these approaches to transform AKI prediction and management, ultimately leading to better patient outcomes and healthcare decision-making. The outcomes of this review are expected to be highly valuable for developing future prediction models in the field of AKI.

The rest of the paper is organized as follows. Section Acute kidney injury describes the proposed methodology, including the review process and selection of relevant studies. Section AKI management presents the results by discussing the descriptive analysis of studies and the analysis of findings. Section Risk prediction models in AKI discusses the findings and generates insights for healthcare professionals on clinically implementing AKI prediction models for improved patient safety, experience, and clinical effectiveness. Further, we present the study limitations and opportunities for future research in Sect. Risk prediction models in AKI. Finally, we present the conclusion of the study in Sect. Methods.

Methods

This systematic review was designed per the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework 2020 [39]. PRISMA framework provides a standard methodology that uses a guideline checklist to ensure transparent and scientifically appropriate systematic reviews. Accordingly, a review protocol was developed to identify the research questions and describe the source of information, search strategy, inclusion and exclusion criteria, data extraction, and analysis procedures. A comprehensive review of the recent literature from January 2012 to January 2023 was undertaken to achieve the aim of this paper and provide healthcare professionals and academics with valuable recommendations on building robust AKI RPMs.

Planning the review

The appropriate selection of databases and keyword identification is essential to ensure a comprehensive and unbiased review [40, 41]. The SLR performed in this study consisted of a search of published articles in three electronic databases: Scopus, PubMed, and EBSCO. The databases were selected based on their recognized impact indices containing scholarly and peer-reviewed medical, systems engineering, and computer science literature. The search was limited to studies published from the year 2012 onwards. The choice for this period is twofold. Firstly, the KDIGO criterion for AKI was published in 2012, and a more standardized definition of AKI has been expected since then [25]. Before that period, there was no standard definition for AKI, therefore, AKI models would vary. Secondly, the healthcare industry has evolved significantly in the last decade, and RPMs developed before 2012 are unlikely to be generalizable to the current settings [42].

Further, after the electronic search, the bibliographies of relevant studies were screened manually to identify additional eligible studies. The search terms and strategy were broadened to identify as many qualifying studies as possible. Search keywords were developed using synonyms/phrases related to AKI, Prediction, and AI, ML tools and combined with Boolean expressions "AND" and "OR," as shown in Table 1.

Table 1 Search string used to obtain articles

SCOPUS, PubMed, and EBSCO		
"Acute Kidney Injury"	**OR**	"Acute Renal Injury" or "Acute Kidney Failure" or "Acute Renal Failure" or "AKI" or "ARI" or "AKF" or "ARF"
AND		
"Prediction"	**OR**	"Predictive" or "Predict" or "Management"
AND		
"Artificial Intelligence"	**OR**	"AI" or "Machine Learning" or "ML" or "Logistics Regression" or "Linear Regression" or "Random Forest" or "Decision Tree" or "Neural Network" or "Deep Learning" or "Bayesian Belief Network" or "Support Vector" or "SVM" or "Bayes Network" or "K-Nearest Neighbor" or "KNN" or "K-means" or "Decision Support" or "Decision Tool"

Table 2 Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
(i) An empirical study	(i) Proceedings of congresses, conference papers, books, book chapters, dissertations, and thesis studies
(ii) Included the adult population	(ii) Review papers
(iii) Written in the English language	(iii) Written in other languages than English
(iv) Available in full text	(iv) Including the children population
(v) Reporting AKI predictive model based on KDIGO definition	(v) Involving patients undergoing kidney transplantation
(vi) Published between 2012 and 2022	(vi) Predicting mortality (in hospital or overall)

To select and include only relevant studies that reported the development of AKI RPMs or externally validated at least one RPM while ensuring viable and unbiased results: the inclusion and exclusion criteria are defined in Table 2.

Conducting the review

The selection of articles to review was carried out in three rounds. The first round consisted of an initial search of the literature included in the EBSCO ($n=87$), PubMed ($n=77$), and SCOPUS ($n=59$) electronic databases with a total of 223 studies. Afterward, Microsoft Excel software was used to exclude duplicate records, yielding 65 studies. Five studies were excluded based on the exclusion criterion (ii) being a review paper resulting in ($n=60$) studies, whereas seven studies were irrelevant to the research objectives, and one study was excluded based on exclusion criteria (iv); therefore, they were excluded as well resulting in a total ($n=52$) studies.

In the second round, titles and abstracts of all included studies were screened, and potentially relevant studies were assessed for eligibility. Fifteen studies were excluded

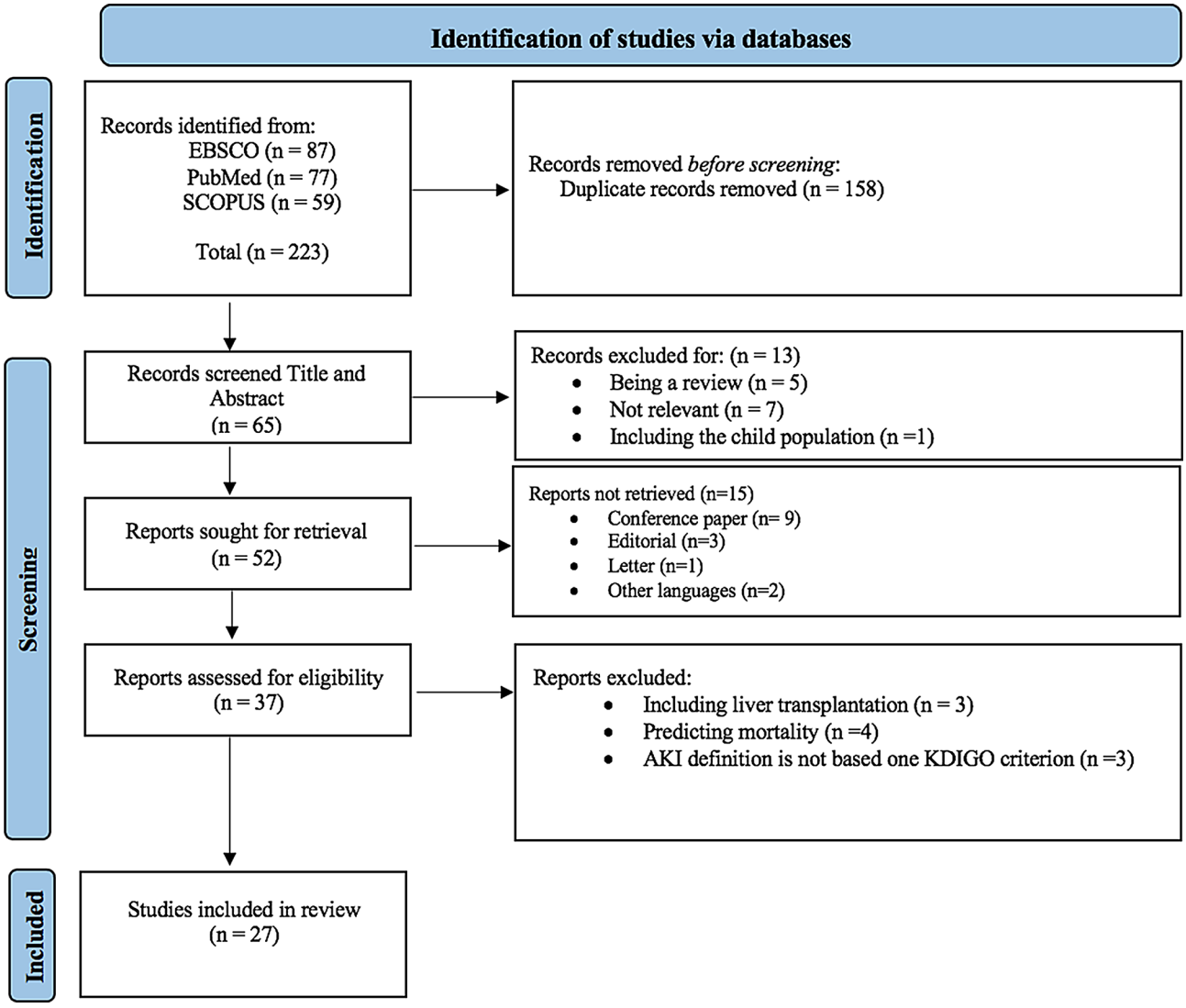

Fig. 1 Describes the literature selection process following the PRISMA 2020 guidelines

based on exclusion criteria (i) and (iii), resulting in a total of (n=37). In the third round, the full texts of all eligible publications were reviewed independently and thoroughly for relevance to the research objectives. Lastly, fourteen studies were excluded from the review based on exclusion criteria (v) and (vi), and the final review included (n=27) studies. Screening of references of the included studies did not reveal additional publications to include. The described search process is summarized in the PRISMA flow diagram in Fig. 1.

Data extraction and quality assessment

A standardized data extraction form was developed and modified as necessary, using a spreadsheet to collect data on existing AKI RPMs characteristics. Items extracted from each study were grouped in five categories namely study-, data-, technology-, modelling- and implementation-related characteristics, as presented in Table 3. Since the literature review aims to identify the risk factors contributing to the development of AKI, the clinical variables are also be reported.

Table 3 Details of extracted items in AKI risk prediction literature

Category	Items
Study-related Characteristics	Year, region, study type, objective
Data-related Related	period, data source, sample size, mean age (years), gender, method of internal validation
Technology-related Related	Algorithms utilized
Modelling-related Related	number of all predictors, predictor selection method, number of significant predictors, AKI indicator, predicted outcome
Implementation-related Related	discrimination AUC, calibration, clinical usefulness, presence of external validation

The final review sample included 27 relevant articles with their related characteristics. The fit between the research methodology and question was assessed to ensure the quality of the studies included in the sample.

Results

Descriptive analysis of the review papers

The 27 reviewed articles were published in 23 different journals. Figure 2 illustrates the distribution of academic journals across a range of categories. Medical journals hold the largest proportion of journals, comprising 70%. The Engineering and Sciences category, accounting for 17% of the journals, demonstrates the presence of publications dedicated to exploring advancements in scientific and engineering disciplines. Noteworthy is the Digital Health and Internet Research category, which accounts for 9% of the journals, signifying an increasing emphasis on the convergence of technology and healthcare within academic research.

Figure 3 shows the number of publications per year. The first study was in 2018, and the number of publications increased significantly afterward. Specifically, 25.9% of the studies were published by the end of 2020, while the majority (74.1%) were published between 2021 and 2022. This increase confirms the importance, interest, and awareness of implementing RPMs for AKI development.

As illustrated in Fig. 4, it is apparent that the sample articles came from various countries worldwide. The countries with the highest number of RPM studies were China (44.44%), followed by the USA and Taiwan (14.8%).

As shown in Figs. 5 and 22 tools were applied in the 27 articles reviewed. Most prediction models utilized are extreme gradient boosting (86.36%), followed by logistics regression (81.82%), and random forest (72.23%). It is also worth noting that support vector machine (45.45%), decision trees (40.91%), and light gradient boosting machine (31.82%) were among the commonly utilized tools in the prediction of AKI. This chart highlights the prominence of specific ML algorithms in predicting AKI within the reviewed literature, providing insights into the preferred approaches in this field of research.

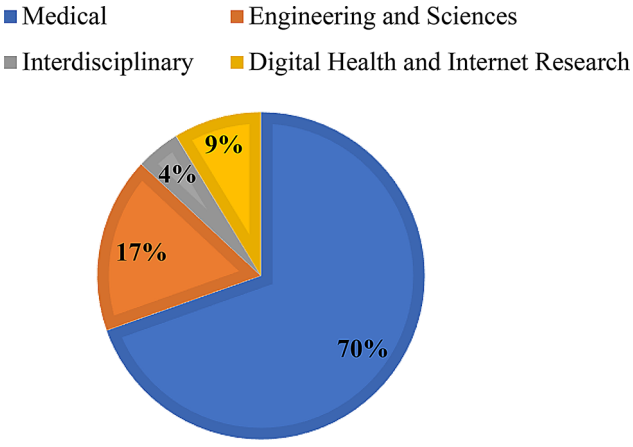

Fig. 2 Distribution of journals across various categories

Study-related characteristic

The characteristics of the studies conducted between 2012 and 2022 are summarized in Table 4. These studies were conducted in different regions with retrospective or prospective study designs.

The objectives of these studies varied, ranging from developing and validating ML models for AKI prediction in specific patient populations to comparing the performance of ML algorithms with traditional statistical models. For instance, one study from Canada focused on developing and validating a preoperative ML model for

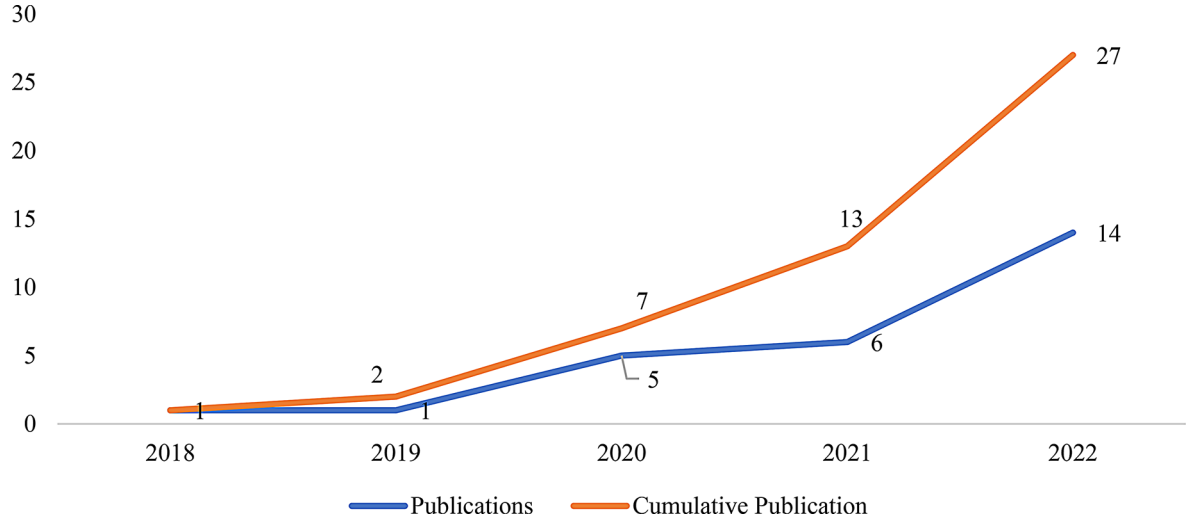

Fig. 3 Number of publications and cumulative publications per year

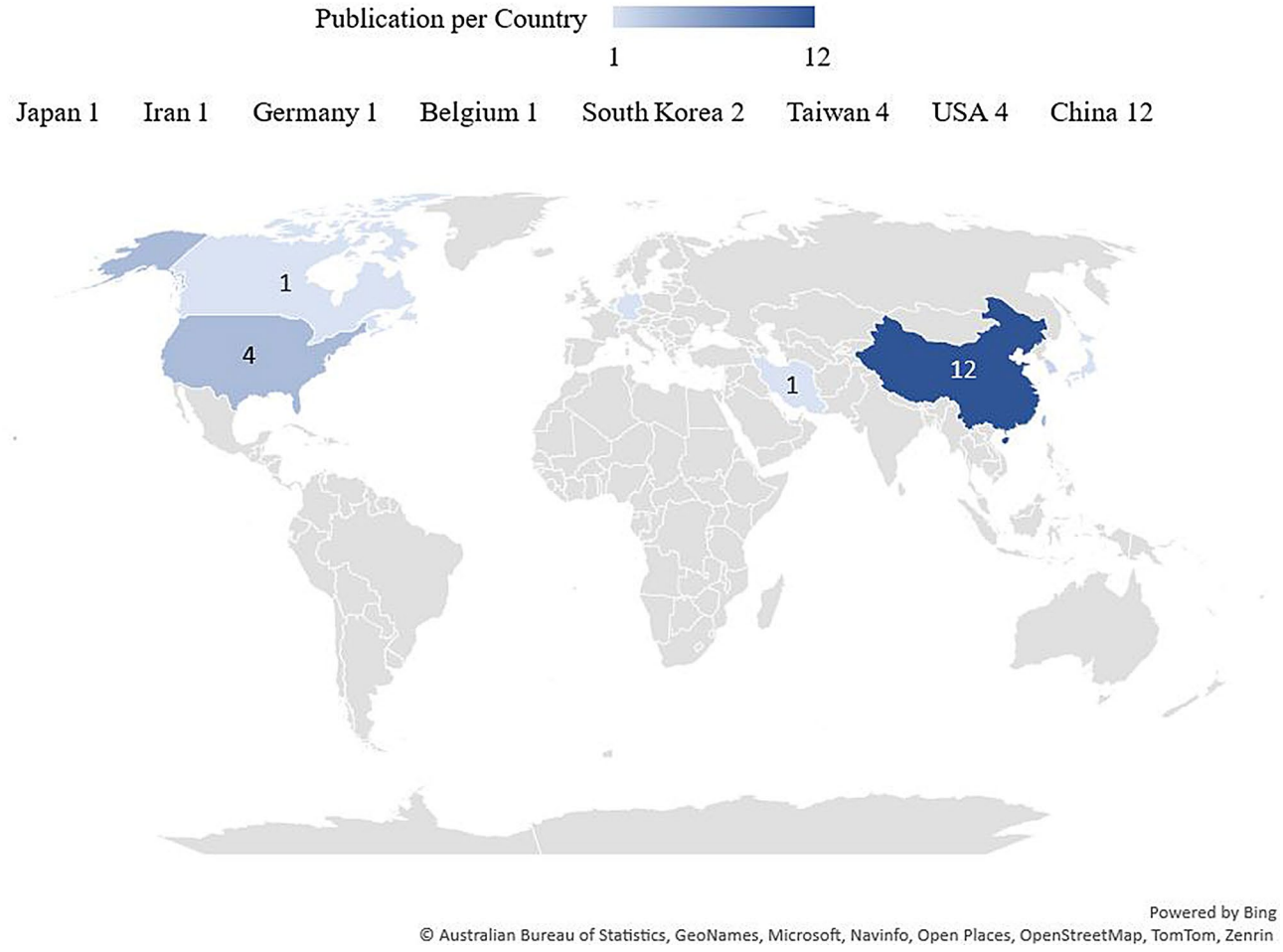

Fig. 4 Number of publications per country

predicting AKI in cardiac surgery patients [43], while a study from China aimed to construct a clinical prediction model for postoperative AKI in patients with acute aortic dissection [52].

Other studies explored the use of ML algorithms to predict AKI risk in patients with different medical conditions, such as sepsis [44, 47, 56], cognitive heart failure [43, 48, 49, 65], acute pancreatitis [53], and femoral neck fractures [50]. Some studies also investigated the feasibility and performance of deep learning algorithms and attention-based temporal neural network approaches for AKI prediction using EHR data [57, 59, 63].

Data-related charactertistics

The study cohort size ranged from 135 to 234,867 patients, depending on the data and study design. Ten studies were conducted in a single-center setting, four were conducted in multiple centers, and the remaining ten used publicly available health data, including the Medical Information Mart for Intensive Care (MIMIC-IV and MIMIC-III) and the electronic ICU Collaborative Research Database (eICU-CRD), which contain data from multiple centers. Additionally, most studies, twenty-four, used retrospective data, two used prospective data, and one used both. Table 5 summarizes the data-related characteristics, including the data collection period, healthcare setting, data source, sample size, mean patient age, gender distribution, and method for internal validation.

The RPMs analyzed in the studies focused on specific patient populations or clinical contexts, such as predicting AKI after cardiac surgery or in patients with sepsis. Seven articles described RPMs for AKI in ICU patients. Six articles covered AKI in patients following major surgery, and three articles addressed sepsis-related patients. While one article focused on a prediction model following major non-cardiac surgery, the other twelve studies focused on other complications. Only one study described a predictive model for a general hospital population [64]. Across many studies, the link between surgical procedures and AKI risk in different hospital settings was explored, and the findings consistently showed that

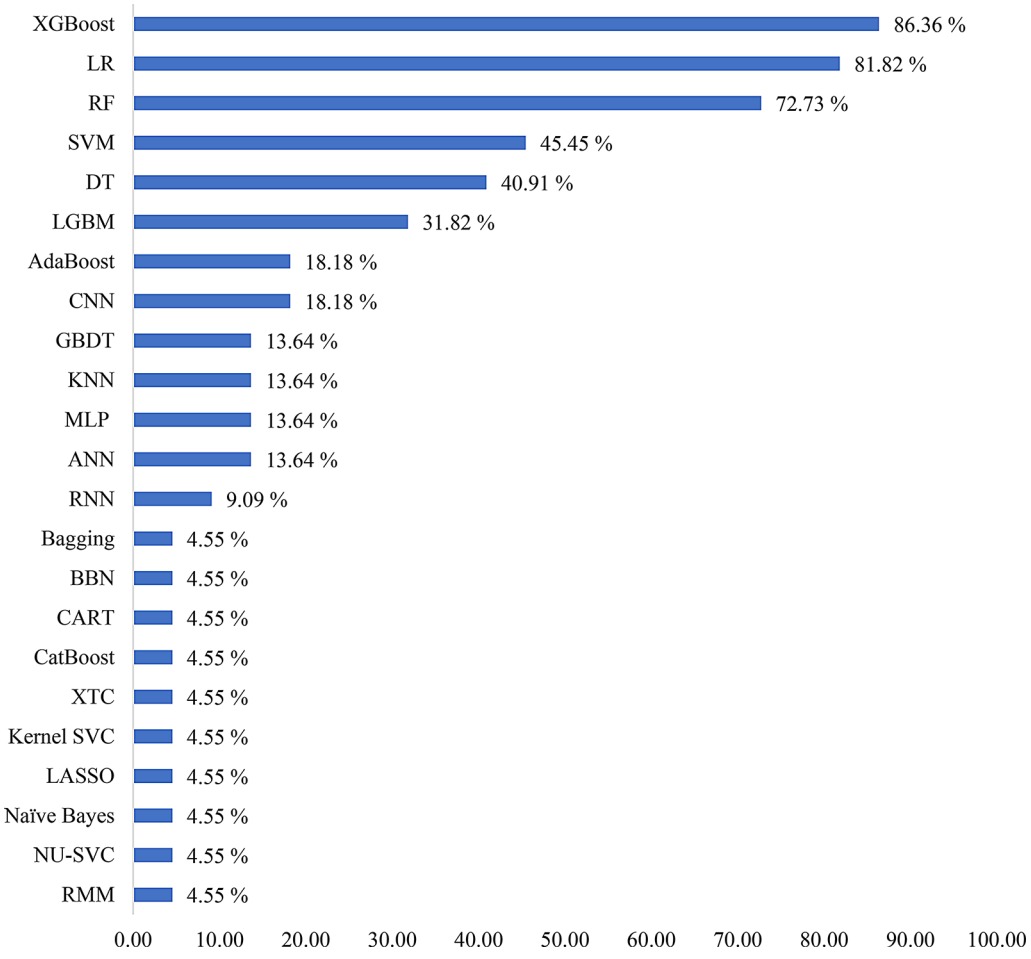

Fig. 5 Percentage of ML algorithms implemented in AKI RPMs. Extreme Gradient Boosting (XGBoost), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), Light Gradient Boosting Machine (LGBM), Adaptive Boosting (AdaBoost), Convolutional Neural Network (CNN), Gradient Boosting Decision Tree (GBDT), K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Bootstrap Aggregation (Bagging), Bayesian Belief Network (BBN), Classification and Regression Trees (CART), Categorical Boosting (CatBoost), Extra Trees Classifier (XTC), Kernel Support Vector Classifier (Kernel SVC), Least Absolute Shrinkage and Selection Operator (LASSO), Naïve Bayes Classifier (Naïve Bayes), Nu-Support Vector Classification (NU-SVC), and Random Mixture Model (RMM)

AKI is a common and significant complication following surgeries. For example, factors such as preoperative heart rate, blood cell counts, albumin, and left ventricular ejection fraction were identified as strong independent predictors of AKI in cardiac surgeries [43, 48, 49, 65]. Additionally, heart failure was strongly associated with AKI development in patients who underwent any surgery, as suggested by multiple studies [50, 53, 63].

Figure 6 indicates the distribution of the settings in which the AKI models were developed or evaluated. The largest portion of the chart, accounting for 41% of the models, corresponds to postoperative care. This significant focus on postoperative care reflects the importance of predicting AKI in patients after surgical procedures, highlighting the need for early detection and intervention in this population. The second-largest portion of the chart, comprising 18% of the models, represents AKI models developed or evaluated in an ICU setting indicating the attention given to critically ill patients. Acute disease patients constitute 15% of the models, emphasizing the relevance of predicting AKI in various acute medical conditions such as heart disease, pancreatic and cerebrovascular. This category stresses the need for accurate risk stratification and early intervention in patients with acute diseases to prevent kidney injury. Both sepsis patients and AKI models in inpatient wards accounted for 11% of each distribution. Interestingly, only one AKI model, accounting for 2% of the total distribution, was specifically developed for patients undergoing CT scans. This focus highlights the risk of AKI associated with CT imaging. Although CT scans are important diagnostic tools, they carry certain risks. One such risk is contrast-induced AKI, which can occur in vulnerable individuals. To reduce the risk of contrast-induced AKI, several preventive measures can be taken. Patients at higher risk, such as those with pre-existing

Table 4 Characteristics of AKI models studies in the literature

Paper	Year	Region	Study Type	Objective
[43]	2022	Canada	Retrospective	Create and internally validate a preoperative ML model to predict AKI associated with cardiac surgery and compare its performance with parametric statistical models.
[44]	2022	China	Retrospective	Establish and validate predictive models for AKI in critically ill sepsis patients using novel ML algorithms.
[45]	2022	USA	Retrospective	Investigate the feasibility of using ML techniques to predict AKI in a Japanese percutaneous coronary intervention database with fewer variables than the NCDR-AKI risk model.
[46]	2022	China	Retrospective	Construct a clinical prediction model for postoperative acute kidney AKI in patients with Type A and Type B acute aortic dissection using ML techniques.
[47]	2022	China	Retrospective	Develop an ensemble ML model to predict early sepsis-associated AKI.
[48]	2022	China	Prospective	Employ a ML algorithm and various techniques for feature selection and model training to enhance the predictive power for cardiac surgery-associated AKI outcomes compared to statistical analysis.
[49]	2022	Iran	Retrospective/Prospective	Develop a predictive tool for cardiac surgery-associated AKI by applying various ML approaches to determine the most effective technique.
[50]	2022	China	Retrospective	Build ML models to predict the risk of AKI in patients with femoral neck fractures.
[51]	2022	China	Retrospective	Establish an ML-based prediction model for early identification of AKI in patients with cognitive heart failure.
[52]	2022	China	Retrospective	Establish a predictive model for acute renal failure occurrence in patients after aortic arch surgery using ML techniques and compare its performance with traditional logistic regression models and other scoring systems.
[53]	2022	China	Retrospective	Develop an AKI risk prediction model for patients with acute pancreatitis.
[54]	2022	China	Retrospective	Determine the incidence and risk factors of AKI and AKI-associated hippocampal damage in patients who have undergone cardiovascular surgery.
[55]	2022	China	Retrospective	Develop and validate an ML model for predicting the risk of AKI in critical care patients with acute cerebrovascular disease.
[56]	2022	Taiwan	Retrospective	Utilize an AI-based ML approach to forecast future risks of rehospitalization with AKI.
[57]	2021	South Korea	Retrospective	Compare the predictive performance of ML algorithms with conventional logistic regression models in assessing the risk of postoperative AKI.
[58]	2021	South Korea	Retrospective	Explore the effectiveness of a deep learning algorithm in anticipating the risk of contrast-induced AKI compared to other ML and logistic regression models in patients undergoing computed tomography.
[37]	2021	Japan	Retrospective	Propose a system for AKI prediction using one-dimensional convolutional neural networks, along with real-time calculation of the probability of developing AKI and visualization of the rationale behind prediction.
[59]	2021	China	Retrospective	Employ an attention-based temporal neural network approach for AKI prediction and analyze the impact of medical features extracted from HER data prior to AKI diagnosis.
[60]	2021	China	Retrospective	Utilize ML methods to identify key indicators and predict AKI in order to provide decision support for clinical staff, requiring measurement of only a small number of indicators.
[61]	2021	USA	Retrospective	Assess the predictive capability of a convolutional neural network-based ML algorithm using patient EHR data to predict AKI stage 2 or 3 up to 48 h in advance of its onset.
[62]	2020	USA	Retrospective	Develop an AKI prediction model and evaluate its transferability across multiple health systems, as well as propose a method to predict the adaptability of AI models in hospitals.
[63]	2020	Germany	Retrospective	Create a deep learning-based algorithm capable of predicting postoperative AKI prior to the manifestation of symptoms and complications.
[64]	2020	Taiwan	Retrospective	Develop, validate, and establish a scoring system using different ML techniques to enable early recognition and intervention of contrast-associated AKI, aiming to prevent permanent kidney damage.
[65]	2020	Taiwan	Retrospective	Employ an AI-based ML approach, driven by perioperative data, to predict contrast-associated AKI.
[66]	2020	China	Retrospective	Establish predictive models for AKI in patients with acute pancreatitis using ML techniques during hospitalization.
[67]	2019	Belgium	Prospective	Prospectively evaluate the performance of the AKI predictor model in predicting AKI-23 within the first week of ICU stay and compare it with predictions made by ICU physicians.
[68]	2018	USA	Retrospective	Evaluate a ML algorithm for early detection and prediction of AKI up to 72 h before its onset, facilitating early clinical intervention.

Table 5 Data-related items

Paper	Period	Setting	Data Source	Sample Size	Mean Age (Years)	Gender, Male (%)	Method of Internal Validation
[43]	2009–2015	Cardiac Surgery	Single Center	6,522	65.8	72.00%	10-CV
[44]	2001–2012	Sepsis	MIMIC-III	3,176	66	55.30%	10-CV
[45]	2008–2019	Cardiovascular	JCD-KiCS	19,222	68.4	-	5-CV
[46]	2019–2021	Cardiovascular	Single Center	283	51.98	80.9%	Bootstrap
[47]	2001–2012	Sepsis	MIMIC-IV	21,038	67	57.6%	CV
[48]	2019–2020	Cardiac Surgery	Single Center	135	-	-	5-CV
[49]	2019–2021	Cardiac Surgery	Multiple Centers	1,620	58.09	60.28	10-CV
[50]	2008–2019	Femoral Neck Fracture	MIMIC-IV	1,596	-	-	10-CV
[51]	2001–2012	Cardiovascular	MIMIC-IV	8,580	-	-	-
[52]	2015–2019	Cardiovascular	Multiple Centers	1,637	-	77.20%	10-CV
[53]	2016–2021	Acute Pancreatitis	Multiple Centers	424	-	62.50%	-
[54]	2019–2020	Cardiovascular Surgery	Single Center	227			5-CV
[55]	2001–2012	Acute Cerebrovascular	MIMIC-III	2,935	66.8	51.10%	5-CV
[56]	2008–2018	Sepsis	Single Center	23,761	76.4	36.00%	-
[57]	2003–2017	Renal Cell Carcinoma	Single Center	4,104	56	69.60%	10-CV
[58]	2007–2019	Intravenous Contrast Media	Single Center	14,185	67.5	77.20%	10-CV
[37]	2014–2015	ICU	eICU-CRD	-	-	-	10-CV
[59]	2001–2012	ICU	MIMIC-III	46,520	-	-	-
[60]	2001–2012	ICU	MIMIC-III	2,018	68.19	49.55%	10-CV
[61]	2008–2012	ICU	MIMIC-III	6,821	60–69	53.29%	10-CV
[62]	2010–2018	ICU	Single Center	153,821	56–65	50.60%	-
[63]	2012–2018	Cardiothoracic Surgery	Single Center	13,895	-	-	5-CV
[64]	2010–2017	General Population	Single Center	234,867	-	-	-
[65]	2016–2018	Cardiac Surgery	Single Center	671	63	67.60%	-
[66]	2014–2019	Acute Pancreatitis	Single Center	334	-	-	-
[67]	2018	ICU	Single Center	252	62.5	61.50%	-
[68]	2001–2017	ICU	Multiple Centers	286,797	-	-	-

MIMIC (Medical Information Mart for Intensive Care), JCD-KiCS (Japanese Cohort of Dialysis and Kidney Injury Critical Study), and eICU-CRD (electronic ICU Collaborative Research Database

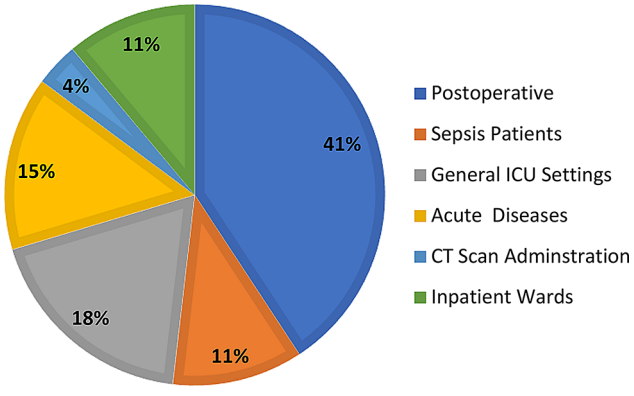

Fig. 6 AKI models settings

kidney conditions or other comorbidities, may be recommended alternative imaging methods that do not use contrast agents.

Modelling-related characteristics

Generally, the definition for predicting AKI was consistent across all studies, following the recent consensus definition of AKI based on KDIGO guidelines as per the inclusion criteria (v) [4]. However, the assessment timeframe varied significantly, depending on the clinical context and specific variables used in the models. Some studies predicted AKI within hours of hospital admission, while others did so after surgery, with timing based on when the surgery occurred. Only one study predicted AKI within a longer timeframe of seven days. This variability in timing highlights the importance of considering clinical settings when developing an RPM and highlights the need to standardize AKI prediction timing. Additionally, the definition of AKI varied depending on the diagnostic criteria. All studies diagnosed AKI based on SCr, except for five studies that included urine output (UO) criteria. All studies used the admission SCr level to represent the patient baseline.

As shown in Table 6, the number of predictors used in the models varied across studies depending on the clinical settings and patient population, ranging from 12 to 134 predictors. The number of significant predictors identified in the models also varied, with studies

Table 6 Modelling-related items

Paper	Number of all Predictors	Predictor Selection Method	No. of Significant Predictors	AKI Indicator	Predicted Outcome
[43]	30	Backward selection	6	SCr	AKI
[44]	36	Boruta algorithm	7	SCr	AKI
[45]	12	LASSO LR, SHAP	7	SCr	AKI
[46]	46 42	SHAP	4	SCr	AKI
[47]	38	Correlation Analysis	17	SCr	AKI Stage 1,2,3
[48]	34	RF LASSO, XGB, SVM	11	SCr and UO	AKI
[49]	44	SHAP, LIME	4	SCr	AKI Staging 1,2,3 for day 1 and 7
[50]	86	SHAP	20	SCr	AKI within 24, 48, 72 h
[51]	58	RFE, SHAP	18	SCr and UO	AKI
[52]	134	SHAP	15	SCr	AKI
[53]	30	Correlation Analysis	10	SCr and UO	AKI Staging 1,2,3
[54]		MLP	7	SCr	AKI
[55]	23	XGB	10	SCr	AKI Staging 1,2,3
[56]	84	SHAP	7	SCr	AKI
[57]	23	Ranking Analysis	5	SCr	Postoperative AKI
[58]	31	SHAP	19	SCr	AKI
[37]		-	84	SCr	AKI Staging 1,2
[59]	54	-	54	SCr	AKI Staging 1,2,3
[60]	47	SFS	17	SCr	AKI
[61]	-	-	-	SCr, UO	AKI Staging 2,3
[62]	13	SHAP	10	SCr	AKI Staging 1,2,3
[63]	96	-	-	SCr	AKI Staging 2,3
[64]	47	XGB, LASSO	10	SCr	AKI
[65]	87	SHAP	20	SCr	AKI
[66]	23	-	5	SCr	AKI Staging 2,3
[67]	-	-	13	SCr	AKI Staging 2,3
[68]	-	-	-	SCr	AKI

LASSO (Least Absolute Shrinkage and Selection Operator), SHAP (SHapley Additive exPlanations), RF (Random Forest), SVM (Support Vector Machine), LIME (Local Interpretable Model-Agnostic Explanations), RFE (Recursive Feature Elimination), XGB (XGBoost), SFS (Sequential Feature Selection) and MLP (Multilayer Perceptron)

reporting between 4 and 42 significant predictors. SCr was the most commonly included predictor, appearing in the majority of studies. Other predictors, such as UO, were also included in some models. The primary predicted outcome was AKI, with some studies further stratifying AKI into different stages. These stages included AKI Stage 1, 2, and 3, as well as AKI incidence within specific timeframes, such as 24, 48, or 72 h.

Most studies used all available predictors to create their models, while others used feature selection techniques to identify key predictors. Three studies utilized all available variables to develop full models, while the remaining twenty-four applied feature selection to identify key predictors and then designed compact models. Among the latter group, ten studies employed the SHapley Additive exPlanations (SHAP) method to assess the importance of individual predictors. This method helps determine which patient factors are most significant for predicting AKI risk, guiding interventions and treatment plans for at-risk patients. Other feature selection methods included backward variable selection, the Boruta algorithm, LASSO logistic regression, Recursive Feature Elimination, and correlation analysis.

For all the RPMs analyzed, the most frequently identified predictors were renal variables such as baseline SCr, bicarbonate, estimated GFR, BUN, and UO, as well as non-renal variables such as age, gender, body mass index (BMI), diabetes, liver failure, mechanical ventilation, medication usage, heart failure, use of vasopressors, diuretics, hemoglobin, white blood cell (WBC) count, blood pressure, hypertension, and hypotension.

Technology and implementation-related characteristics

Table 7 provides the results of different algorithms used in the development. A range of algorithms were employed, including random forest, logistics regression, gradient boosting variants, neural networks and decision trees.

Each algorithm was evaluated based on its discrimination ability, represented by the Area under the receiver operating curve (AUROC). The studies had different objectives; some aimed to create predictive models that

Table 7 Technology and implementation-related items

Paper	Algorithm	Discrimination AUC	Calibration	Clinical Usefulness	External Validation
[43]	RF + LR	0.75	H-L Test	-	No
[44]	XGB	0.82	Calibration Plots	DCA	No
[45]	LGDM	0.77	Calibration Plots	-	No
[46]	RF	0.76	Calibration Plots, Brier Score	DCA	No
[47]	XGB	0.77–0.79	-	-	Yes
[48]	RF	0.86	Barier Score	DCA	No
[49]	RF	0.81	-	-	No
[50]	LGBM	0.93	-	-	NO
[51]	LGBM	0.80	Calibration Plots	-	Yes
[52]	XGB	0.82	Brier Scre (0.087)	-	Yes
[53]	RF	0.91	Calibration Plots	DCA CIC	Yes
[54]	RF	0.82	-	-	No
[55]	XGBoost	0.88	Calibration Plots	-	Yes
[56]	LGBM	0.82	-	-	No
[57]	LGBM	0.81	H-L Test	DCA	No
[58]	RNN	0.76	-	-	No
[37]	CNN	0.84	-	-	No
[59]	NN	0.83	-	-	Yes
[60]	XGBoost	0.93	-	-	Yes
[61]	CNN	0.86	-	-	No
[62]	DS-GBT	0.78	H-L Test	-	Yes
[63]	RNN	0.89	Calibeation Plots H-L Test, Brier Score	-	No
[64]	XGB	0.79	-	-	No
[65]	RF + XGB	0.84	-	-	No
[66]	XGB	0.92	-	-	No
[67]	RF	0.80	Calibration Plot	-	Yes
[68]	DT	0.87	-	-	No

RF (Random Forest), LR (Logistic Regression), XGB (Extreme Gradient Boosting), LGDM (Logistic Gradient Descent Machine), LGBM (Light Gradient Boosting Machine), RNN (Recurrent Neural Network), CNN (Convolutional Neural Network), DS-GBT (Dual-Stage Gradient Boosting Tree), DT (Decision Tree) and H-L Test (Hosmer–Lemeshow Test)

were simple and easy for clinicians to interpret, while others prioritized achieving high predictive accuracy. The performance metrics for these models included discrimination, calibration, clinical usefulness, and the presence of external validation. All models underwent internal validation, and some studies also conducted external validation and assessed clinical usefulness.

Discrimination was measured using the AUROC, with a median AUROC of 0.753 across the studies (ranging from 0.590 to 0.930). For studies that performed external validation, which indicates whether the model was validated on an independent dataset, the median AUROC for the development model was 0.772 (ranging from 0.592 to 0.929), and for external validation, the median AUROC was 0.78 (ranging from 0.69 to 0.9142). Calibration plots were utilized in several studies to assess the calibration of the models, with some reporting good calibration outcomes, while others did not provide specific calibration results. The Hosmer-Lemeshow (H-L) test, a statistical test for goodness-of-fit, was employed in a few studies to evaluate calibration. The p-values from the H-L test indicated the level of agreement between predicted and observed outcomes. Several studies also assessed clinical usefulness using metrics such as decision curve analysis (DCA) and the Brier score. DCA measures the net benefit of the model across a range of threshold probabilities, while the Brier score evaluates the accuracy of predicted probabilities. These metrics provide insights into the practical utility of the models in clinical decision-making. The presence of external validation varied across the studies. While some models were externally validated, others were not, leading to potentially overly optimistic results from internal validation based on the same cohort. This highlights the need for further validation of AKI prediction models.

Notably, in situations involving the application of multiple ML algorithms, only a single model using the highest-performing algorithm has been reported. Therefore, each study was regarded as having one model. The method of dealing with missing data was defined in a few studies, with some using bootstrap resampling or other techniques. In some cases, imputation methods were not

used as the advanced boosting ML method can handle missing values automatically. For other models, missing values of continuous variables were imputed using median values, while categorical variables were imputed using mode values. Additionally, some studies used a k-nearest neighbors' approach to fill in missing values.

Discussion

Key findings

The studies demonstrate that leveraging EHR data has significantly increased interest in AI and ML-based risk prediction models for healthcare. These models assist physicians in anticipating future events, particularly in AKI prediction, where precise predictions are essential for proactive monitoring and intervention, thus reducing AKI incidence and severity while improving quality of care. A review of 27 articles was conducted, focusing on RPMs for AKI.

The RPMs analyzed show considerable diversity in clinical settings, patient characteristics, cohort regions, and statistical methods. Most models focus on AKI in hospitalized patients rather than community-acquired AKI, reflecting that most studies have been conducted in developed regions where AKI is primarily hospital-acquired. However, developing countries with limited healthcare access may have higher rates of community-acquired AKI, limiting these models' generalizability and ability to inform prevention strategies. A recent study estimated AKI incidence at 4.3% among all hospital admissions [69]. Yet, this figure remains an underestimate of the true occurrence of community AKI due to the non-referral of patients to hospitals. For this reason, prevention strategies for AKI should consider both hospital and community-acquired cases to be effective. Furthermore, many models are tailored to specific patient groups and settings, such as cardiac surgeries or sepsis patients. The goal of developing AKI RPMs for surgery-related cases is to identify risk factors during or before procedures that pose a high risk for AKI. This allows clinicians to weigh the benefits and risks of procedures, determine when to monitor kidney function, and take preventive measures.

The reported incidences of AKI were inconsistent and differed significantly, depending on the definition of AKI, the population studied, and the clinical context in which it is detected. Therefore, to use AKI RPMs in clinical practice, their cross-site transportability, or ability to perform consistently in different healthcare settings/populations, must be validated. A cross-site transportable model should be able to provide reliable predictions regardless of the location or context in which it is applied and is thereby generalizable. This is an important consideration in many fields, including healthcare, where predictive models are often developed and validated in one setting but must be applied in different settings with different patient populations, data collection methods, geographical locations, periods, and clinical workflows. Failure of a prediction model to transport well across new settings indicates that the model cannot be readily implemented in clinical practice for new patients. Thereby, ensuring cross-site transportability is important to ensure the model is ready for use in local patients and avoid potential errors that may arise from using a model in inappropriate settings. Only one study assessed and predicted the transportability of AKI prediction models and found that their model could accurately predict AKI at all external sites, demonstrating cross-site transportability [62]. Their study suggests that using ML models for AKI prediction can be generalized across different healthcare systems, leading to better patient outcomes. However, their study also found that the model's performance slightly decreased when applied to other multiple centers. More research is needed to fully comprehend the transportability of AKI RPMs and develop models that perform well in diverse healthcare settings.

The geographical distribution of the studies reviewed highlights the concentration of AKI prediction research in certain regions, particularly in East Asia and North America. Most studies were conducted in China, the USA, and other developed regions like Japan, Germany, and South Korea. This regional concentration may introduce biases that affect the applicability of AI models in more diverse or underrepresented healthcare environments, particularly those with limited resources. Recognizing this concentration helps emphasize the need for further research in underrepresented regions, where differences in healthcare systems, data availability, and resource allocation could significantly affect the clinical applicability of AI models. By understanding these regional contexts, the broader utility and adaptability of AI-based AKI prediction models can be better assessed, which is crucial for ensuring their effectiveness across various healthcare settings.

AKI often occurs in patients with common risk factors, particularly in specific medical settings. Most studies (74%) provided the rationale for selecting candidate predictors, indicating that model development was primarily data-driven with good specification of known clinically important factors. There was substantial variability across models in the variables included as candidate predictors. From over 50 predictive risk factors identified across the 27 studies, 20 factors were used consistently in more than one study. We identified 20 common variables frequently included in prediction models, and choosing appropriate candidate predictors is essential for deriving accurate RPMs. These models generally use various demographic, clinical, and laboratory factors to predict the likelihood of AKI in different hospital settings.

Among these commonly included predictors, SCr was the most prevalent. SCr is widely utilized as it is an established biomarker for evaluating kidney function and detecting renal impairment, making it an important predictor for AKI. However, the use of SCr as both a predictor and a diagnostic criterion requires careful consideration to avoid conflating its roles in early identification and diagnosis. In addition, UO was included in some models, although its incorporation was limited by data availability. UO is an important marker for early detection, often providing a more immediate indication of reduced kidney function compared to SCr [70, 71]. One study showed that adding UO criteria can detect AKI in patients 11 h earlier than SCr criteria alone [70]. Nevertheless, we believe it is not ideal to include both SCr and UO as independent variables in the model because they are highly correlated [72]. Typically, elevated SCr levels indicate impaired kidney function, which is often associated with a decrease in UO, suggesting an inverse relationship between the two variables [73]. Therefore, including both variables in the model may not significantly improve its predictive performance, as they essentially provide overlapping information about kidney function.

Other renal-related variables, such as BUN, eGFR, and bicarbonate, were also commonly used. BUN and eGFR are well-known indicators of renal function, while bicarbonate helps assess acid-base balance, which is particularly relevant for patients at risk of kidney injury [72]. Collectively, these renal variables provide a comprehensive understanding of kidney function, which can support early intervention efforts to prevent AKI progression.

Beyond renal-specific indicators, non-renal variables were also identified as significant contributors to AKI risk, highlighting the systemic nature of the condition. Age was frequently included, as older individuals are more prone to AKI due to age-related renal changes and increased comorbidity burdens, including cardiovascular disease and diabetes. High BMI was another common predictor, as obesity is linked to inflammation and metabolic stress, both of which can contribute to kidney damage [74]. Gender differences were also evident, with male patients generally found to be at higher risk of AKI compared to female patients, likely due to physiological and hormonal factors.

Several comorbid conditions were consistently included as predictors, including diabetes, hypertension, cardiac diseases and liver failure, all of which can affect vascular health and increase the likelihood of renal injury [75]. Factors associated with critical illness, such as mechanical ventilation, use of vasopressors, and diuretics, were also found to be significant predictors, as these factors reflect severe conditions that can compromise kidney perfusion and increase the risk of injury due to nephrotoxic exposure. Additionally, Sepsis was identified as a common risk factor for AKI due to the kidney's sensitivity to hypoperfusion, mechanical ventilation, and excessive fluid resuscitation [47, 56]. When the body responds to Sepsis, it triggers an inflammatory response that activates innate immunity and releases proinflammatory substances that often lead to AKI [76].

Lastly, laboratory markers such as hemoglobin levels, WBC count, and blood pressure (including both hypertension and hypotension) were frequently included in prediction models. These markers provide insights into systemic health, the presence of inflammatory or hypoxic conditions, and cardiovascular function, all of which can affect renal outcomes. Including these demographics, clinical, and laboratory variables in prediction models reflects the complexity of AKI risk and highlight the need to adopt a comprehensive approach that considers both kidney-specific and systemic factors when developing predictive models.

Some RPMs initially use all available predictors to build a comprehensive model, followed by feature selection to create a reduced model containing only the most significant predictors. While a full model includes all potential predictors relevant to AKI prediction, this often results in a complex model that may be challenging for clinicians to use in practice. Therefore, reducing the number of predictors helps balance model performance and usability, ensuring the model remains clinically applicable. For example, a study by Li et al. [52] developed a full model using 134 predictors. However, gathering such a large number of clinical variables in real-life practice can be impractical and time-consuming. To address this, the authors developed a reduced model with only 15 predictors, which performed similarly to the full model but was easier to apply in clinical settings with less computational power required. Reduced models often offer comparable performance to full models while being easier to interpret and use, which supports findings from previous studies. Future RPMs for AKI should focus on readily and routinely available factors to provide timely interventions and enhance clinical usability. The SHAP method was frequently utilized to illustrate the influence of the key indicators on the risk of AKI. It was observed that certain variables have a greater impact on the model's ability to predict AKI. Our findings suggest that ML tools can accurately predict AKI by considering the complex interactive relationship between several important variables.

Internal validation was conducted by randomly partitioning the dataset, but this approach is limited since the derivation and validation datasets share substantial similarities. In contrast, other studies used more robust techniques such as bootstrapping and cross-validation (CV), which are preferred due to their use of the entire dataset during model development and their reduced likelihood

of overfitting [77]. Although internal validation demonstrated good performance, these results may be overly optimistic. This issue is commonly encountered when predictive models, particularly RPMs, are applied to different samples from those used for model development, leading to a decline in prediction accuracy.

External validation was rarely performed in the reviewed studies, and when done, it revealed decreased model accuracy, limiting the generalizability of these models across diverse medical centers and settings [78, 79]. The lack of external validation could be due to challenges in accessing external datasets, including patient privacy concerns, data-sharing restrictions, and logistical issues. Differences in data collection methods, patient populations, and healthcare environments also complicate external validation without significant adjustments to the original model. To enhance confidence in the applicability of these models across various hospital settings and geographic regions, future studies should prioritize external validation to thoroughly assess their efficacy.

AKI prediction models are typically developed using retrospective data of individuals at risk for AKI, incorporating known baseline and follow-up kidney function information. While retrospective data can provide valuable insights into past events, it may face limitations such as missing data and selection bias. Appropriate statistical methods should be used to address these biases and improve the generalizability of the models. Recent studies have shown that combining retrospective and prospective data can improve the quality of the information available [63]. Nevertheless, further prospective studies are essential to validate the findings derived from retrospective analyses. Another limitation of existing models is the use of single-center data, which may not be representative of larger populations and could lead to poor model performance when applied to other cohorts [80]. To overcome these limitations, incorporating data from multiple centers can increase sample size and improve model generalizability, as demonstrated in several studies [49, 52, 53, 68].

Overall, gradient boosting models were the most effective in predicting AKI, with XGB exhibiting superior performance compared to other ML tools. XGB is decision tree ensemble method that uses a gradient framework to enhance the effectiveness of decision trees. It can be applied to both classification and regression problems and is known for handling complex datasets with numerous predictors and missing data. Due to its strong performance, XGB is increasingly used to predict adverse clinical outcomes, and its results have been shown to outperform models such as K-nearest neighbors, logistic regression, decision trees, random forest, and neural networks [81, 82]. Based on these findings, extreme gradient boosting is an efficient and reliable ML algorithm for building prediction models. In contrast, traditional logistics regression may show lower accuracy because it relies on handling linear combinations of variables. This can lead to oversimplification of complex non-linear relationships in clinical RPMs. LR is also affected by multicollinearity, which may result in the omission of important associations and limit the model's predictive capacity.

Good calibration was reported for most models. Besides assessing discrimination, calibration is an important measure of a risk prediction model's validity, as it indicates the level of agreement between predicted and observed probabilities of developing AKI, which is essential for effectively communicating risk to patients and care providers. Furthermore, most studies did not demonstrate clinical usefulness, which limits the ability to recommend one model over another. Although the potential of ML tools and RPMs to enhance AKI predictions is promising, further research is required to validate and optimize these models for diverse patient populations.

Our review emphasizes the need for further research to develop reliable and accurate risk prediction models for AKI that can be applied to the general population. Although one model developed by Hsu et al. using a large dataset and common risk factors may serve as a foundation for such efforts, further refinements are required for broader applicability [64]. Importantly, RPMs are not necessarily designed to directly enhance decision-making for AKI management; rather, they assist clinicians in selecting the most appropriate model for their specific clinical settings and intended use.

Gaps and potential solutions for AKI management

ML has shown significant promise in healthcare, particularly in predicting AKI, improving patient outcomes and reducing healthcare costs. Developing accurate and reliable RPMs is essential to identify individuals at high risk of AKI and enable early intervention. However, as illustrated in Fig. 7, several obstacles affect the accurate prediction of AKI incidences. These barriers fall into five categories: study-related, data-related, technology-related, modeling-related, and implementation-related gaps.

Study-related gaps include insufficient data on community-acquired AKI, which leads to underestimating its occurrence and limits prevention strategies. Variations in the time frames used for reporting AKI incidence also make it difficult to assess the true burden of the condition. Additionally, using single-center data limits model performance across diverse populations. To address these gaps, systems should be established to link data from multiple healthcare facilities, providing a more comprehensive view of AKI cases, including both community-acquired and hospital-acquired AKI. Defining consistent

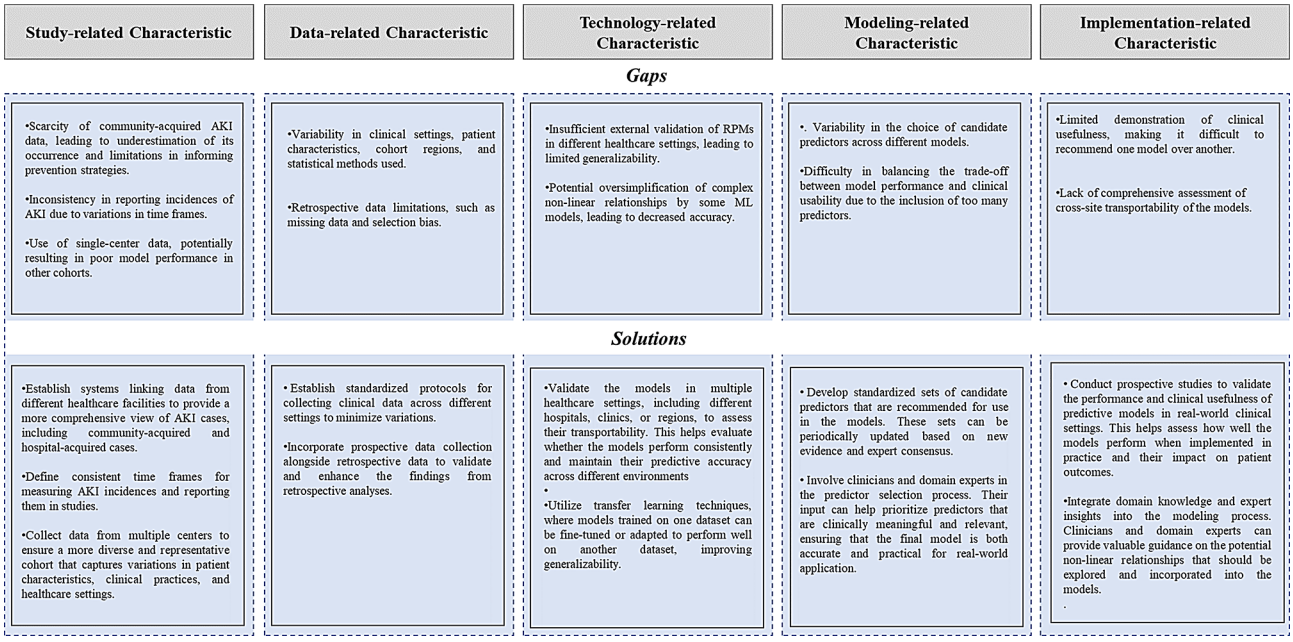

Fig. 7 Gaps and potential solutions for AKI prediction models

time frames for measuring and reporting AKI incidence is also necessary for comparability across studies.

Data-related challenges encompass variability in clinical settings, patient characteristics, cohort regions, and statistical methods. Retrospective data issues, such as missing information and selection bias, also affect model robustness. Standardized protocols for data collection across different settings should be established to minimize variations and enhance generalizability. Incorporating prospective data collection can further validate and improve findings from retrospective analyses.

Technology-related gaps are primarily due to insufficient external validation of RPMs in diverse healthcare settings, which limits generalizability. Some ML models also oversimplify complex non-linear relationships, reducing predictive accuracy. Validating models across multiple healthcare settings, including hospitals and clinics in different regions, is essential for assessing their transportability. Transfer learning techniques can further improve generalizability by adapting models trained on one dataset to perform well on others.

Modeling-related challenges include variability in the choice of candidate predictors across different models and the difficulty in balancing the trade-off between model performance and clinical usability due to the inclusion of too many predictors. Developing standardized sets of candidate predictors that are recommended for use in models can help overcome these issues. These sets can be periodically updated based on new evidence and expert consensus. Involving clinicians and domain experts in the predictor selection process is also important to prioritize predictors that are clinically meaningful and relevant, ensuring that the final model is both accurate and practical for real-world application.

Implementation-related gaps involve the limited demonstration of clinical usefulness, making it challenging to recommend one model over another, and the lack of comprehensive assessment of cross-site transportability of the models. Conducting prospective studies to validate the performance and clinical usefulness of predictive models in real-world clinical settings helps assess how well the models perform when implemented in practice and their impact on patient outcomes. Integrating domain knowledge and expert insights into the modeling process can provide valuable guidance on the potential non-linear relationships that should be explored and incorporated into the models.

Given the methodological limitations identified in previous systematic reviews of AKI RPMs and their external validation, large datasets collected from multiple centers using consensus AKI criteria are necessary to derive and validate accurate AKI outcome prediction models. By addressing these challenges and implementing the proposed solutions, the accuracy, generalizability, clinical utility, and value of AKI risk prediction models can be significantly enhanced, ultimately leading to improved patient outcomes.

Practical implications for clinicians

While the development of RPMs using ML has demonstrated significant potential in identifying AKI risks, implementing these models in clinical practice presents several challenges for clinicians. One key challenge is the complexity of certain models, which makes them difficult

to interpret and use in time-sensitive situations. Clinicians are often unfamiliar with how ML algorithms arrive at predictions, which may lead to aversion in adopting these tools without a clear understanding of their functionality. This highlights the importance of explainable AI (XAI) tools that can provide transparent reasoning for predictions and help improve trust and usability among clinicians [62]. By offering insights into why certain features contribute to the predicted outcomes, XAI can help bridge the gap between AI and clinical decision-making, making models more acceptable for real-world application.

Another important aspect for practical clinical adoption is the balance between model complexity and usability [83]. Several of the studies reviewed used reduced models that prioritized fewer but clinically relevant predictors, maintaining predictive power while simplifying practical implementation. Reduced models are beneficial because they maintain similar accuracy levels to full models but require fewer clinical variables, which makes them easier to apply in a real-world healthcare setting. This approach can save valuable time, reduce computational requirements, and enhance the clinician's experience when applying the model in a busy clinical environment. Thus, future efforts should continue focusing on developing RPMs that prioritize clinical relevance and simplicity to maximize their practical usability without sacrificing predictive performance.

Limitation and future research

Our research presents a comprehensive review of AKI RPMs, based on the KDIGO consensus published in 2012. This review identifies risk variables, examines prediction models, and compiles a list for future validation. It highlights the utility of AKI risk factors in accurately predicting AKI. However, practical application requires focusing on clinically available variables, as not all identified predictors are suitable for routine use in clinical practice.

Despite the strengths of this review, our findings must be interpreted considering several limitations. Most studies were conducted in developed countries, limiting broader applicability to low-resource settings. Including data from underrepresented regions and focusing on community-acquired AKI will improve generalizability. First, the reviewed studies often lacked external validation, limiting the generalizability of the models across diverse healthcare settings. The absence of external validation means that the consistency and reliability of model performance in different clinical environments remain uncertain. Another limitation is the reliance on retrospective data in most studies, which can introduce biases such as selection bias and missing data, thereby affecting the robustness of the findings. Prospective data collection would provide a more reliable basis for model development and validation by reducing the impact of these biases.

Additionally, the heterogeneity of the included studies—with variations in study design, populations, data sources, and predictors—posed a challenge for comparison and synthesis. This lack of uniformity prevented a quantitative meta-analysis and necessitated a more descriptive approach, limiting the ability to draw strong statistical conclusions.

In summary, while our review offers valuable insights into the development and validation of AKI RPMs, the findings emphasize the need for further research that focuses on external validation, prospective data, standardized methodologies, and simpler, clinically feasible models. Addressing these limitations will improve the reliability and generalizability of AKI prediction models, ultimately enhancing their practical application in diverse healthcare settings.

Conclusion

AKI is a major global health issue, with serious consequences including longer hospital stays and rising healthcare costs. The use of ML-based RPMs presents an opportunity to improve patient care by identifying those at high risk of developing AKI. However, the wide range of patient types and healthcare settings makes it difficult to establish a standard approach for managing these high-risk individuals. Although some models show potential, their impact on patient outcomes has not yet been fully demonstrated, and there is no agreement on the most effective models due to existing limitations. To address these challenges, it is essential to gather detailed clinical data in a consistent format, covering relevant risk factors for AKI. Developing reliable and practical models requires efforts to include clinical details that current medical records may not easily capture. Standardizing data collection across healthcare facilities could help create broadly applicable AKI risk prediction models. Additional research is needed to enhance the precision and relevance of these models for a variety of medical conditions. Testing model reliability in multiple settings is essential to ensure their effectiveness in real-world use and their true impact on patient outcomes. Progress in AKI risk prediction modeling will contribute significantly to improving patient care and outcomes. Therefore, efforts to develop improved and widely applicable AKI risk prediction models should remain a key focus of medical research.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-024-02758-y.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
D.A. and M.S. wrote the main manuscript text and prepared all figures and tables. M.O. and S.A. validated the results and revised the manuscript. All authors reviewed and approved the manuscript.

Funding
Not applicable.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 20 March 2024 / Accepted: 8 November 2024
Published online: 14 November 2024

References

1. Hilton R. Acute renal failure. BMJ [Internet]. 2006 Oct 12 [cited 2023 Feb 13];333(7572):786–90. https://www.bmj.com/content/333/7572/786
2. Bellomo R, Kellum JA, Ronco C. Acute kidney injury. Lancet. 2012;380(9843):756–66.
3. Susantitaphong P, Cruz DN, Cerda J, Abulfaraj M, Alqahtani F, Koulouridis I, et al. World incidence of AKI: a meta-analysis. Clin J Am Soc Nephrol. 2013;8(9):1482–93.
4. Chawla LS, Bellomo R, Bihorac A, Goldstein SL, Siew ED, Bagshaw SM, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 workgroup. Nat Rev Nephrol. 2017;13(4):241–57.
5. Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW. Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. J Am Soc Nephrol. 2005;16(11):3365–70.
6. Li PKT, Burdmann EA, Mehta RL, World Kidney Day Steering Committee. 2013. Acute kidney injury: global health alert. Kidney Int. 2013;83(3):372–6.
7. Acker PV, Biesen WV, Nagler EV, Koobasi M, Veys N, Vanmassenhove J. Risk prediction models for acute kidney injury in adults: An overview of systematic reviews. PLoS ONE [Internet]. 2021 Apr 1 [cited 2024 Oct 23];16(4):e0248899. https://pmc.ncbi.nlm.nih.gov/articles/PMC8016311/
8. Zhou H, Liu L, Zhao Q, Jin X, Peng Z, Wang W et al. Machine learning for the prediction of all-cause mortality in patients with sepsis-associated acute kidney injury during hospitalization. Front Immunol [Internet]. 2023 Apr 3 [cited 2024 Oct 23];14. https://www.frontiersin.org/journals/immunology/articles/https://doi.org/10.3389/fimmu.2023.1140755/full
9. Li X, Wu R, Zhao W, Shi R, Zhu Y, Wang Z, et al. Machine learning algorithm to predict mortality in critically ill patients with sepsis-associated acute kidney injury. Sci Rep. 2023;13(1):5223.
10. Grams ME, Rabb H. The distant organ effects of acute kidney injury. Kidney International [Internet]. 2012 May 2 [cited 2023 Feb 13];81(10):942–8. https://www.kidney-international.org/article/S0085-2538(15)55192-2/fulltext
11. Haredasht FN, Vanhoutte L, Vens C, Pottel H, Viaene L, De Corte W. Validated risk prediction models for outcomes of acute kidney injury: a systematic review. BMC Nephrology [Internet]. 2023 May 9 [cited 2024 Oct 23];24(1):133. https://doi.org/10.1186/s12882-023-03150-0
12. Nateghi Haredasht F, Viaene L, Pottel H, De Corte W, Vens C. Predicting outcomes of acute kidney injury in critically ill patients using machine learning. Sci Rep [Internet]. 2023 Jun 18 [cited 2024 Oct 23];13(1):9864. https://www.nature.com/articles/s41598-023-36782-1
13. Fischer MJ, Brimhall BB, Parikh CR. Uncomplicated Acute Renal Failure and Post-Hospital Care: A Not So Uncomplicated Illness. Am J Nephrol [Internet]. 2008 Apr [cited 2023 Feb 13];28(3):523–30. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2786012/
14. Vijayan A, Abdel-Rahman EM, Liu KD, Goldstein SL, Agarwal A, Okusa MD et al. Recovery after Critical Illness and Acute Kidney Injury. Clinical Journal of the American Society of Nephrology: CJASN [Internet]. 2021 Oct [cited 2024 Oct 23];16(10):1601. https://pmc.ncbi.nlm.nih.gov/articles/PMC8499012/
15. Igiraneza G, Dusabejambo V, Finklestein FO, Rastegar A. Challenges in the Recognition and Management of Acute Kidney Injury by Hospitals in Resource-Limited Settings. Kidney International Reports [Internet]. 2020 Jul 1 [cited 2024 Oct 23];5(7):991–9. https://www.sciencedirect.com/science/article/pii/S2468024920311748
16. Silver SA, Harel Z, McArthur E, Nash DM, Acedillo R, Kitchlu A et al. 30-Day Readmissions After an Acute Kidney Injury Hospitalization. The American Journal of Medicine [Internet]. 2017 Feb 1 [cited 2023 Feb 13];130(2):163–172.e4. https://www.amjmed.com/article/S0002-9343(16)31023-3/fulltext
17. Siew ED, Peterson JF, Eden SK, Hung AM, Speroff T, Ikizler TA, et al. Outpatient nephrology referral rates after acute kidney injury. J Am Soc Nephrol. 2012;23(2):305–12.
18. Lai CF, Wu VC, Huang TM, Yeh YC, Wang KC, Han YY, et al. Kidney function decline after a non-dialysis-requiring acute kidney injury is associated with higher long-term mortality in critically ill survivors. Crit Care. 2012;16(4):R123.
19. Rewa O, Bagshaw SM. Acute kidney injury-epidemiology, outcomes and economics. Nat Rev Nephrol. 2014;10(4):193–207.
20. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. Lancet. 2019;394(10212):1949–64.
21. Ali T, Khan I, Simpson W, Prescott G, Townend J, Smith W, et al. Incidence and outcomes in acute kidney injury: a comprehensive population-based study. J Am Soc Nephrol. 2007;18(4):1292–8.
22. Mehta RL, McDonald B, Gabbai F, Pahl M, Farkas A, Pascual MTA, et al. Nephrology consultation in acute renal failure: does timing matter? Am J Med. 2002;113(6):456–61.
23. Yang L, Xing G, Wang L, Wu Y, Li S, Xu G, et al. Acute kidney injury in China: a cross-sectional survey. Lancet. 2015;386(10002):1465–71.
24. Mehta RL, Cerdá J, Burdmann EA, Tonelli M, García-García G, Jha V, et al. International Society of Nephrology's 0by25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. Lancet. 2015;385(9987):2616–43.
25. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. Nephron Clin Pract. 2012;120(4):c179–184.
26. Finlay S, Bray B, Lewington A, Hunter-Rowe C, Banerjee A, Atkinson J et al. Identification of risk factors associated with acute kidney injury in patients admitted to acute medical units. Clin Med (Lond) [Internet]. 2013 Jun [cited 2023 Feb 28];13(3):233–8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5922664/
27. Janes H, Pepe MS, Gu W. Assessing the Value of Risk Predictions Using Risk Stratification Tables. Ann Intern Med [Internet]. 2008 Nov 18 [cited 2023 Feb 15];149(10):751–60. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3091826/
28. Kok JN, Koronacki J, de Mantaras RL, Matwin S, Mladenič D, Skowron A, Machine Learning, editors. ECML 2007: 18th European Conference on Machine Learning, Warsaw, Poland, September 17–21, 2007. Proceedings [Internet]. Berlin, Heidelberg: Springer; 2007 [cited 2023 Feb 17]. (Lecture Notes in Computer Science; vol. 4701). http://link.springer.com/https://doi.org/10.1007/978-3-540-74958-5
29. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in Medicine [Internet]. 2001 Aug [cited 2023 Feb 17];23(1):89–109. https://linkinghub.elsevier.com/retrieve/pii/S093336570100077X
30. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. N Engl J Med [Internet]. 2016 Sep 29 [cited 2023 Feb 17];375(13):1216–9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5070532/
31. Li J, Tran M, Siwabessy J. Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. PLoS One [Internet]. 2016 Feb 18 [cited 2023 Feb 21];11(2):e0149089. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4758710/

32. Chen JH, Asch SM. Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations. N Engl J Med [Internet]. 2017 Jun 29 [cited 2023 Feb 17];376(26):2507–9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5953825/
33. Huang H, Liu Y, Wu M, Gao Y, Yu X. Development and validation of a risk stratification model for predicting the mortality of acute kidney injury in critical care patients. Ann Transl Med [Internet]. 2021 Feb [cited 2023 Mar 3];9(4):323. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7944298/
34. Park S, Lee H. Acute kidney injury prediction models: current concepts and future strategies. Curr Opin Nephrol Hypertens. 2019;28(6):552–9.
35. Josephs SA, Thakar CV. Perioperative risk assessment, prevention, and treatment of acute kidney injury. Int Anesthesiol Clin. 2009;47(4):89–105.
36. Lee HC, Yoon HK, Nam K, Cho YJ, Kim TK, Kim WH, et al. Derivation and Validation of Machine Learning Approaches to predict acute kidney Injury after Cardiac surgery. J Clin Med. 2018;7(10):322.
37. 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.
38. Holmes J, Allen N, Roberts G, Geen J, Williams JD, Phillips AO et al. Acute kidney injury electronic alerts in primary care - findings from a large population cohort. QJM: An International Journal of Medicine [Internet]. 2017 Sep 1 [cited 2023 Feb 14];110(9):577–82. https://doi.org/10.1093/qjmed/hcx080
39. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ [Internet]. 2021 Mar 29 [cited 2023 Feb 17];372:n71. https://www.bmj.com/content/372/bmj.n71
40. McDonagh M, Peterson K, Raina P, Chang S, Shekelle P. Avoiding Bias in Selecting Studies. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 [cited 2023 Feb 20]. (AHRQ Methods for Effective Health Care). http://www.ncbi.nlm.nih.gov/books/NBK126701/
41. Ahn E, Kang H. Introduction to systematic review and meta-analysis. Korean J Anesthesiol [Internet]. 2018 Apr [cited 2023 Feb 20];71(2):103–12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5903119/
42. Kim KK, Michelman JE. An Examination of Factors for the Strategic Use of Information Systems in the Healthcare Industry. MIS Quarterly [Internet]. 1990 [cited 2023 Feb 20];14(2):201–15. https://www.jstor.org/stable/248778
43. Petrosyan Y, Mesana TG, Sun LY. Prediction of acute kidney injury risk after cardiac surgery: using a hybrid machine learning algorithm. BMC Medical Informatics and Decision Making [Internet]. 2022 May 18 [cited 2023 Feb 20];22(1):137. https://doi.org/10.1186/s12911-022-01859-w
44. Yue S, Li S, Huang X, Liu J, Hou X, Zhao Y et al. Machine learning for the prediction of acute kidney injury in patients with sepsis. Journal of Translational Medicine [Internet]. 2022 May 13 [cited 2023 Feb 20];20(1):215. https://doi.org/10.1186/s12967-022-03364-0
45. Kuno T, Mikami T, Sahashi Y, Numasawa Y, Suzuki M, Noma S et al. Machine learning prediction model of acute kidney injury after percutaneous coronary intervention. Sci Rep [Internet]. 2022 Jan 14 [cited 2023 Feb 20];12(1):749. https://www.nature.com/articles/s41598-021-04372-8
46. Xinsai L, Zhengye W, Xuan H, Xueqian C, Kai P, Sisi C, et al. Prediction model of acute kidney injury after different types of acute aortic dissection based on machine learning. Front Cardiovasc Med. 2022;9:984772.
47. Zhang L, Wang Z, Zhou Z, Li S, Huang T, Yin H, et al. Developing an ensemble machine learning model for early prediction of sepsis-associated acute kidney injury. iScience. 2022;25(9):104932.
48. Xue X, Liu Z, Xue T, Chen W, Chen X. Machine learning for the prediction of acute kidney injury in patients after cardiac surgery. Frontiers in Surgery [Internet]. 2022 [cited 2023 Jun 12];9. https://www.frontiersin.org/articles/https://doi.org/10.3389/fsurg.2022.946610
49. Ejmalian A, Aghaei A, Nabavi S, Abedzadeh Darabad M, Tajbakhsh A, Abin AA et al. Prediction of Acute Kidney Injury After Cardiac Surgery Using Interpretable Machine Learning. Anesth Pain Med [Internet]. 2022 Sep 28 [cited 2023 Mar 16];12(4). https://brieflands.com/articles/aapm-127140.html
50. Liu J, Xu L, Zhu E, Han C, Ai Z. Prediction of acute kidney injury in patients with femoral neck fracture utilizing machine learning. Front Surg [Internet]. 2022 Jul 26 [cited 2023 Feb 20];9:928750. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9360500/
51. Peng X, Li L, Wang X, Zhang H. A machine learning-based prediction model for acute kidney Injury in patients with congestive heart failure. Front Cardiovasc Med. 2022;9:842873.
52. Li J, Gong M, Joshi Y, Sun L, Huang L, Fan R et al. Machine Learning Prediction Model for Acute Renal Failure After Acute Aortic Syndrome Surgery. Frontiers in Medicine [Internet]. 2022 [cited 2023 Feb 20];8. https://www.frontiersin.org/articles/https://doi.org/10.3389/fmed.2021.728521
53. 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;15:5061–72.
54. Xue X, Chen W, Chen X. A Novel Radiomics-based machine learning Framework for Prediction of Acute kidney Injury-related delirium in patients who underwent Cardiovascular surgery. Comput Math Methods Med. 2022;2022:4242069.
55. 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 [Internet]. [cited 2023 Jun 12];44(1):43–53. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8856083/
56. Ou SM, Lee KH, Tsai MT, Tseng WC, Chu YC, Tarng DC. Artificial Intelligence for Risk Prediction of Rehospitalization with Acute Kidney Injury in Sepsis Survivors. Journal of Personalized Medicine [Internet]. 2022 Jan [cited 2023 Feb 20];12(1):43. https://www.mdpi.com/2075-4426/12/1/43.
57. Lee Y, Ryu J, Kang MW, Seo KH, Kim J, Suh J et al. Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma. Sci Rep [Internet]. 2021 Aug 3 [cited 2023 Feb 20];11(1):15704. https://www.nature.com/articles/s41598-021-95019-1
58. Yun D, Cho S, Kim YC, Kim DK, Oh KH, Joo KW, et al. Use of Deep Learning to predict acute kidney Injury after intravenous contrast Media Administration: Prediction Model Development Study. JMIR Med Inf. 2021;9(10):e27177.
59. Chen Z, Chen M, Sun X, Guo X, Li Q, Huang Y et al. Analysis of the Impact of Medical Features and Risk Prediction of Acute Kidney Injury for Critical Patients Using Temporal Electronic Health Record Data With Attention-Based Neural Network. Frontiers in Medicine [Internet]. 2021 [cited 2023 Feb 20];8. https://www.frontiersin.org/articles/https://doi.org/10.3389/fmed.2021.658665
60. Wang J, Zhu B, Liu P, Jia R, Jia L, Chen W, COMMUNICATIONS & CONTROL [Internet]. Screening Key Indicators for Acute Kidney Injury Prediction Using Machine Learning. INTERNATIONAL JOURNAL OF COMPUTERS. 2021 Apr 12 [cited 2023 Feb 20];16(3). https://univagora.ro/jour/index.php/ijccc/article/view/4180
61. Le S, Allen A, Calvert J, Palevsky PM, Braden G, Patel S, et al. Convolutional Neural Network Model for Intensive Care Unit Acute kidney Injury Prediction. Kidney Int Rep. 2021;6(5):1289–98.
62. Song X, Yu ASL, Kellum JA, Waitman LR, Matheny ME, Simpson SQ et al. Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. Nat Commun [Internet]. 2020 Nov 9 [cited 2023 Feb 20];11(1):5668. https://www.nature.com/articles/s41467-020-19551-w
63. Rank N, Pfahringer B, Kempfert J, Stamm C, Kühne T, Schoenrath F et al. Deep-learning-based real-time prediction of acute kidney injury outperforms human predictive performance. npj Digit Med [Internet]. 2020 Oct 26 [cited 2023 Feb 20];3(1):1–12. https://www.nature.com/articles/s41746-020-00346-8
64. Hsu CN, Liu CL, Tain YL, Kuo CY, Lin YC. Machine learning model for risk prediction of Community-Acquired Acute Kidney Injury Hospitalization from Electronic Health Records: Development and Validation Study. J Med Internet Res. 2020;22(8):e16903.
65. Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. Critical Care [Internet]. 2020 Jul 31 [cited 2023 Feb 20];24(1):478. https://doi.org/10.1186/s13054-020-03179-9
66. Qu C, Gao L, Yu X, qiang, Wei M, Fang G, quan, He J et al. Machine Learning Models of Acute Kidney Injury Prediction in Acute Pancreatitis Patients. Gastroenterology Research and Practice [Internet]. 2020 Sep 29 [cited 2023 Feb 20];2020:e3431290. https://www.hindawi.com/journals/grp/2020/3431290/
67. Flechet M, Falini S, Bonetti C, Güiza F, Schetz M, Van den Berghe G et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. Critical Care [Internet]. 2019 Aug 16 [cited 2023 Feb 20];23(1):282. https://doi.org/10.1186/s13054-019-2563-x
68. Mohamadlou H, Lynn-Palevsky A, Barton C, Chettipally U, Shieh L, Calvert J, 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.
69. Wonnacott A, Meran S, Amphlett B, Talabani B, Phillips A. Epidemiology and outcomes in community-acquired versus hospital-acquired AKI. Clin J Am Soc Nephrol. 2014;9(6):1007–14.

70. Koeze J, Keus F, Dieperink W, van der Horst ICC, Zijlstra JG, van Meurs M. Incidence, timing and outcome of AKI in critically ill patients varies with the definition used and the addition of urine output criteria. BMC Nephrol. 2017;18(1):70.
71. Vanmassenhove J, Steen J, Vansteelandt S, Morzywolek P, Hoste E, Decruyenaere J, et al. The importance of the urinary output criterion for the detection and prognostic meaning of AKI. Sci Rep. 2021;11(1):11089.
72. Pandya D, Nagrajappa AK, Ravi KS. Assessment and Correlation of Urea and Creatinine Levels in Saliva and Serum of Patients with Chronic Kidney Disease, Diabetes and Hypertension– A Research Study. J Clin Diagn Res [Internet]. 2016 Oct [cited 2023 Feb 27];10(10):ZC58–62. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5121806/
73. Engoren M, Maile MD, Heung M, Jewell ES, Vahabzadeh C, Haft JW et al. The Association Between Urine Output, Creatinine Elevation, and Death. The Annals of Thoracic Surgery [Internet]. 2017 Apr [cited 2024 Oct 23];103(4):1229–37. https://linkinghub.elsevier.com/retrieve/pii/S0003497516309353
74. Grams ME, Sang Y, Ballew SH, Gansevoort RT, Kimm H, Kovesdy CP, et al. A Meta-analysis of the Association of Estimated GFR, Albuminuria, Age, Race, and sex with acute kidney Injury. Am J Kidney Dis. 2015;66(4):591–601.
75. Virzì G, Day S, de Cal M, Vescovo G, Ronco C. Heart-kidney crosstalk and role of humoral signaling in critical illness. Crit Care. 2014;18(1):201.
76. White LE, Chaudhary R, Moore LJ, Moore FA, Hassoun HT. Surgical sepsis and organ crosstalk: the role of the kidney. J Surg Res. 2011;167(2):306–15.
77. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453–73.
78. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130(6):515–24.
79. Bleeker S, Moll HA, Steyerberg E, Donders R, Derksen-Lubsen G, Grobbee D, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol. 2003;56:826–32.
80. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Feb 25]. (Statistics for Biology and Health). http://link.springer.com/https://doi.org/10.1007/978-3-030-16399-0
81. Zhu Y, Zhang J, Wang G, Yao R, Ren C, Chen G et al. Machine Learning Prediction Models for Mechanically Ventilated Patients: Analyses of the MIMIC-III Database. Frontiers in Medicine [Internet]. 2021 [cited 2023 Mar 13];8. https://www.frontiersin.org/articles/https://doi.org/10.3389/fmed.2021.662340
82. Liu J, Wu J, Liu S, Li M, Hu K, Li K. Predicting mortality of patients with acute kidney injury in the ICU using XGBoost model. PLoS ONE. 2021;16(2):e0246306.
83. Cutillo CM, Sharma KR, Foschini L, Kundu S, Mackintosh M, Mandl KD et al. Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency. npj Digit Med [Internet]. 2020 Mar 26 [cited 2024 Oct 24];3(1):47. https://www.nature.com/articles/s41746-020-0254-2

Publisher's note