The Open Urology & Nephrology Journal
DOI: 10.2174/011874303X508237261030114818, 2026, 19, e1874303X508237
ISSN: 1874-303X
1

MINI-REVIEW **OPEN ACCESS**

Artificial Intelligence and Machine Learning for Prognostication Across the AKI–AKD–CKD Continuum: Predictive Models, Clinical Outcomes, and the Path Toward Clinical Translation

Manikandan Alagesan[1] and Venkateswaramurthy Nallasamy[1,*]

[1]*Department of Pharmacy Practice, J.K.K. Nattraja College of Pharmacy, Kumarapalayam, India*

Abstract:

Introduction: Acute kidney injury affects about one in five hospitalised adults and substantially increases the risk of chronic kidney disease, yet conventional scoring systems poorly identify patients who will progress. This review critically appraises artificial intelligence and machine learning models for prognostication across the acute kidney injury–acute kidney disease–chronic kidney disease continuum.

Methods: A narrative review of 48 source articles was undertaken, covering persistent kidney injury, renal recovery, dialysis dependence, major adverse kidney events, mortality, and chronic kidney disease. Studies were examined for prediction endpoint, injury aetiology, algorithm, dataset, validation strategy, interpretability, and clinical applicability.

Results: Reported discrimination varied substantially across algorithms, endpoints, populations, datasets, and validation strategies, with several models outperforming conventional comparators in within-study comparisons. Serum creatinine dynamics, blood urea nitrogen, age, vasopressor use, urine output, and measures of comorbidity, including the Charlson Comorbidity Index (CCI), were recurrent predictors, although their importance varied by population and endpoint. Fewer than one-third of studies were externally validated; most were retrospective and at high risk of bias, calibration and fairness were seldom assessed, and no randomised evidence of improved patient outcomes exists.

Discussion: The evidence comprises heterogeneous, non-equivalent prediction tasks rather than a single modelled transition to chronic kidney disease. Cross-study performance comparisons are therefore unreliable, and post-hoc explanation tools cannot replace inherent interpretability, calibration, and prospective validation.

Conclusion: Artificial intelligence shows promise for prognostication after acute kidney injury, but clinical translation requires clearly defined endpoints, aetiology-aware modelling, prospective multicentre validation, fairness assessment, and randomised trials demonstrating patient benefit.

Keywords: Artificial intelligence, Machine learning, Acute kidney injury, Acute kidney disease, Chronic kidney disease, Prognostication, Clinical prediction model.

© 2026 The Author(s). Published by Bentham Open.
This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: https://creativecommons.org/licenses/by/4.0/legalcode. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Address correspondence to this author at the Department of Pharmacy Practice, J.K.K. Nattraja College of Pharmacy, Kumarapalayam, India; Ph: +91 98427 24689; E-mail: nvmurthi@gmail.com

Cite as: Alagesan M, Nallasamy V. Artificial Intelligence and Machine Learning for Prognostication Across the AKI–AKD–CKD Continuum: Predictive Models, Clinical Outcomes, and the Path Toward Clinical Translation. Open Urol Nephrol J, 2026; 19: e1874303X508237. http://dx.doi.org/10.2174/011874303X508237261030114818

Received: May 06, 2026
Revised: September 08, 2026
Accepted: September 23, 2026
Published: November 05, 2026

Send Orders for Reprints to
reprints@benthamscience.net

1. INTRODUCTION

Acute kidney injury (AKI) is a common clinical syndrome that has been defined as the rapid decrease of excretory kidney function. Its global burden has been

conclusively quantified: one in five adults and one in three children have acute kidney injury (AKI) during episodes of hospital care according to KDIGO criteria, as shown by a meta-analysis consisting of 312 studies with almost 49.2 million patients [1]. The associated mortality is significant, with pooled AKI-associated mortality rates of 23.9% in adults and 13.8% in children [1]. In the ICU, the incidence increases to 50–60% of all admissions [2], carrying a heavy burden in terms of length of stay, healthcare costs, and long-term morbidity [2, 3].

However, the long-term consequences of AKI have increasingly been recognised as an important component of its clinical burden. It has also been shown that AKI survivors are at significantly higher risk for developing CKD, with a pooled adjusted hazard ratio of 8.8 for subsequent CKD and 3.1 for end-stage renal disease (ESRD) compared to patients without AKI [4], and experience contemporary rates of CKD incidence reaching as high as 25.8 per 100 person-years [4]. Importantly, patients who achieve seemingly full recovery are still at significantly increased risk of incident CKD long-term, and most are not referred for nephrology follow-up. This finding is both a knowledge gap and an avenue for intervention [5].

The pathophysiological pathways linking AKI with subsequent kidney dysfunction are heterogeneous and depend partly on the initiating insult. Ischaemic/hemodynamic, sepsis-associated, nephrotoxic/drug-induced, and perioperative AKI have overlapping but distinct mechanisms of injury and repair; therefore, the pathways leading to persistent dysfunction and CKD should not be assumed to be identical across etiologies. Shared downstream processes can include epithelial G2/M cell-cycle arrest, cellular senescence, profibrogenic signalling, pericyte activation, capillary rarefaction, and interstitial fibrosis [6, 7]. Epigenetic changes, including DNA methylation and histone modifications, may further influence repair and fibrosis [8]. The ADQI framework nevertheless provides a useful temporal continuum by defining AKD as persistent or ongoing kidney dysfunction between 7 and 90 days after AKI [9].

Conventional risk stratification instruments such as KDIGO staging, SOFA scores, and APACHE-II have limited prognostic accuracy in distinguishing patients at risk of progressing from AKI to CKD [10]. Other novel biomarkers for persistent AKI have been described (*e.g.*, urinary CCL14, NGAL, KIM-1 and [TIMP-2] × [IGFBP7]), but they are limited by cost, availability and challenges in integration into routine practice [11, 12]. In this context, AI and ML have been applied to several distinct prediction tasks after AKI, including persistent AKI/AKD, renal recovery, dialysis-related outcomes, major adverse kidney events, mortality, and longer-term CKD progression [13, 14]. This review summarises and critically appraises evidence from 48 source articles across the AKI-AKD-CKD continuum, with explicit attention to prediction endpoint, AKI context or etiology, algorithm, dataset, validation strategy, interpretability, and clinical applicability.

2. METHODOLOGY

A structured narrative literature search was conducted to identify relevant evidence on artificial intelligence (AI) and machine learning (ML) for prognostication across the acute kidney injury (AKI)–acute kidney disease (AKD)–chronic kidney disease (CKD) continuum. PubMed/MEDLINE, Scopus, and Web of Science were searched for articles published from January 1995 to March 2026, with Google Scholar used as a supplementary source to identify additional relevant publications. Earlier landmark articles, consensus statements, and methodological frameworks were included where necessary through citation tracking.

Search terms were used individually and in combination, including “acute kidney injury,” “acute kidney disease,” “chronic kidney disease,” “AKI-to-CKD progression,” “renal recovery,” “persistent AKI,” “dialysis dependence,” “major adverse kidney events,” “artificial intelligence,” “machine learning,” “deep learning,” “prediction model,” “prognosis,” “XGBoost,” “random forest,” “neural network,” and “clinical prediction.” Peer-reviewed original studies, systematic reviews, consensus documents, and methodological guidance relevant to prediction of persistent AKI/AKD, renal recovery, dialysis-related outcomes, major adverse kidney events, mortality, or subsequent CKD were considered. Relevant studies were critically evaluated according to study population, AKI context or etiology, prediction endpoint, AI/ML algorithm, dataset, validation strategy, model performance, interpretability, and clinical applicability. Owing to substantial heterogeneity in populations, outcome definitions, prediction horizons, and modelling approaches, the evidence was synthesized narratively

3. PATHOPHYSIOLOGICAL BASIS OF THE AKI-AKD-CKD CONTINUUM: THE RATIONALE FOR AI PREDICTION

Information on the biological and physiological mechanisms that contribute to AKI-to-CKD progression is important for assessing AI prediction models, as these mechanisms shape the temporal windows, biomarkers, and clinical variables that computational systems need to exploit. The kidney has a remarkable regenerative capacity after an acute renal insult, primarily through the proliferation of surviving tubular epithelial cells. However, when injury is acute or recurrent and severe, a range of maladaptive processes act to undermine normal repair and promote progression towards CKD [6].

Key among these is epithelial cell-cycle arrest in the G2/M phase, which promotes a profibrogenic phenotype typified by upregulation of TGF-β and connective tissue growth factor [7]. Meanwhile, cellular senescence induces a senescence-associated secretory phenotype (SASP) that sustains chronic inflammation and fibrosis [6]. Along with capillary rarefaction, pericyte detachment and transdifferentiation into myofibroblasts contribute to progressive interstitial fibrosis and vascular loss. Epigenetic regulatory processes that control gene expression patterns crucial to fibrogenesis and repair [8]

modulate these dynamics. This ultimately results in progressive loss of nephron function and increasing fibrosis, which irreversibly reduces glomerular filtration rate (GFR)—the hallmarks of persistent kidney damage or chronic kidney disease (CKD) [6].

AKI is now recognised as existing within a spectrum of acute kidney disease (AKD), which fills the interval between 7 days of AKI and CKD that develops after 90 days [9]. Both AKI and AKD represent a continuum of kidney injury that can ultimately lead to CKD [9, 15]. This continuum encompasses heterogeneous renal patterns, ranging from recovery to injury persistence to progression to chronic disease, and serves as a conceptual basis for risk stratification across all phases of kidney injury and recovery. Importantly, the AKI–AKD–CKD continuum is a temporal framework rather than a single homogeneous disease trajectory. Etiology, severity, recurrence, baseline kidney function, and recovery pattern may influence both biological progression and predictive feature relevance. Accordingly, models predicting persistent AKI/AKD, renal recovery, dialysis dependence, MAKE, mortality, and CKD after 90 days are considered related but non-equivalent prediction tasks in this review. Under this framework, machine learning methodologies have been utilised for specific prediction tasks, such as AKI recovery of renal function [2]. Together, these studies emphasise the expanding role of AI across the AKI–CKD continuum; however, current methods focus primarily on independent prediction tasks rather than an integrated, end-to-end risk prediction system.

4. ARTIFICIAL INTELLIGENCE METHODS APPLIED ACROSS THE AKI- CKD CONTINUUM

4.1. The Shift from Static to Dynamic Prediction

Predictive modelling for AKI has progressed from a static, single-time-point risk assessment paradigm to dynamic, continuously updated models [13, 14]. Using procedure-specific and preoperative features, early strategies determined risk at a fixed point in time, whereas modern models incorporate changing clinical data over the course of hospitalisation and repeatedly recalculate risk. This change is particularly important for longitudinal outcome prediction, where time-varying clinical features may provide more insight into disease pathways than single-timepoint assessments.

The evolution has been facilitated by the increasing accessibility of large-scale electronic health record (EHR) databases. These databases include MIMIC-III and the eICU Collaborative Research Database, which have been frequently used as data sources for AKI prediction studies [16, 17]. Though these datasets are useful, it is concerning that models trained on them may have generalizability issues because such data often come from discrete healthcare environments and very few models are externally validated or are applicable, as highlighted in systematic reviews [17, 18].

4.2. Gradient Boosting Algorithms: XGBoost and LightGBM

Among ML methods applied to AKI-related outcomes, XGBoost and LightGBM are frequently used and have demonstrated strong discrimination in several individual datasets. Their reported performance, however, depends on the study population, outcome definition, input variables, preprocessing, and validation strategy; these studies should therefore not be interpreted as establishing universal superiority of gradient boosting over other algorithms. Selected high-performing AI/ML models across the AKI–AKD–CKD continuum, including their prediction endpoints, reported AUROC values, and principal predictive factors, are summarized visually in Fig. (**1**).

These methods fit a sequence of weak decision trees, where each later tree attempts to correct the mistakes of those before it, and regularisation is used to prevent overfitting. They have advantages, including natural handling of missing data, relatively efficient computation, and the ability to model complex non-linear relationships in heterogeneous clinical datasets.

In a pediatric context, an XGBoost model was created to predict MAKE30 and 90-day adverse outcomes (chronic dialysis and death) in 1,394 children hospitalised with AKI. AUROCs achieved for MAKE30 and 90-day outcomes were 0.810 and 0.851, respectively, which outperformed logistic regression (AUROCs of 0.786 and 0.759) [19]. The tool was implemented as a web-based risk calculator for bedside application [19]. For older inpatients, an XGBoost model with the Boruta feature selection algorithm to predict MAKE30 in 4,266 patients achieved training and internal test AUROCs of 0.868 and 0.823, respectively, while external validation on 11,864 patients from MIMIC-IV yielded an AUROC of 0.744 [20]. The 5 variables found to be most closely related to MAKE30 were vasopressor use, mechanical ventilation at the time of hospitalisation, blood urea nitrogen, red blood cell distribution width, and serum albumin according to SHAP analysis [20].

Analysis of AKD risk and associated mortality in critically ill older adults using the LightGBM algorithm. In a cohort of 3,542 elderly AKI patients from the MIMIC-IV database with external validation at a Chinese hospital, LightGBM attained AUROCs of 0.844, 0.853, and 0.755 for predicting AKD risk, and AUROCs of 0.861, 0.868, and 0.746 for predicting AKD-associated mortality [21]. SHAP analysis indicated that the most important predictors were creatinine on day 3, sepsis, delta blood urea nitrogen, diastolic blood pressure, and heart rate. The models were implemented as web-based applications for clinical use [21].

A multicentre cohort study developed four machine-learning algorithms-logistic regression, random forest, support vector machine, and XGBoost-to predict hospital mortality and major adverse kidney events (MAKE) in 9,587 critically ill patients with incident AKI [10]. A model comprising 71 clinical variables obtained during the initial three days of ICU stay was trained, followed by feature selection based on both model importance and clinical

reasoning, yielding 14-feature prediction models for MAKE and 15-feature prediction models for mortality. Within this multicentre cohort, the ML models showed higher discrimination than the corresponding KDIGO and SOFA comparators; however, these within-study comparisons do not establish that a particular ML algorithm will perform better across different institutions or datasets.

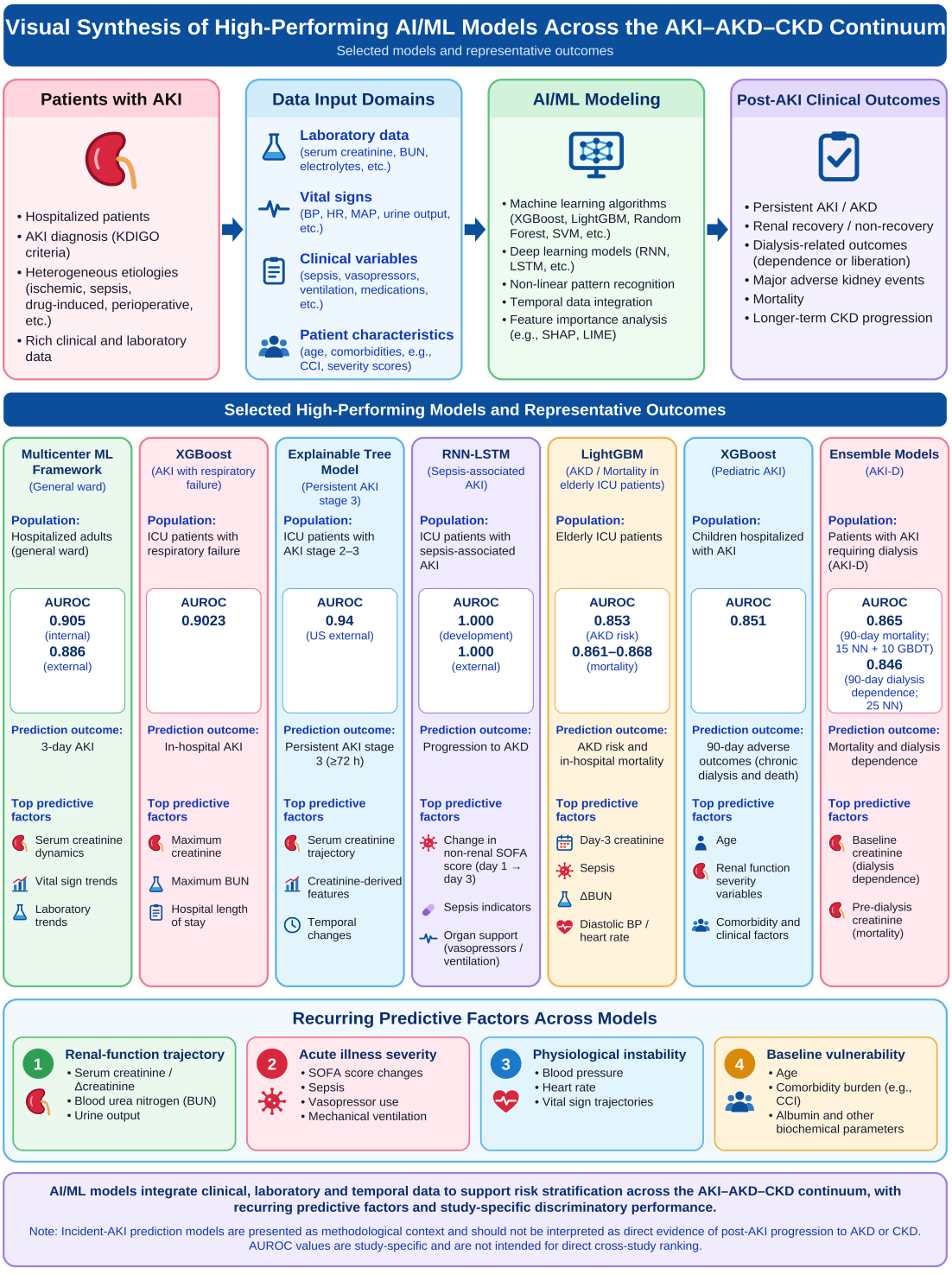

Fig. (1). Visual synthesis of high-performing AI/ML models across the AKI–AKD–CKD continuum.

Gradient-boosting methods have also been applied to incident AKI prediction in specific populations, including ICU patients with respiratory failure and patients with gout [22, 23]. These studies are relevant examples of algorithm use but address AKI onset rather than post-AKI progression and therefore are not considered direct evidence of transition to CKD. In the gout cohort, the reported model also evaluated AKD, which is more directly relevant to the post-AKI continuum [23]. In ICU patients with respiratory failure, XGBoost attained an AUROC of 0.9023 among a sample of 10,780 MIMIC-IV patients, with creatinine maximum as well as hospital length of stay and BUN maximum yielding the most significant predictive weight [22]. LightGBM models achieved AUROCs of 0.815 for AKI and 0.873 for AKD in gout patients, indicating that specific clinical identifiers such as diuretics, serum sodium levels, and urate-lowering therapy agents are particularly important for predicting disease [23].

4.3. Random Forest and Survival Analysis Approaches

Random forest algorithms and their survival-analysis adaptations, including random survival forests, have been evaluated for post-AKI CKD and mortality prediction. In individual cohorts, they have shown useful discrimination, but performance should be interpreted as cohort- and endpoint-specific rather than as evidence of general algorithm superiority. In a prospective observational study of 101 critically ill patients with AKI stage 3, random forest models developed to predict CKD at 3 and 6 months achieved AUPR values of 0.895 and 0.848, respectively, outperforming baseline logistic regression models. The concordance index for survival prediction using RSF and survival XGBoost was approximately 0.825 [24]. After 3 months, CKD had developed in 47 of the 75 surviving patients, and after 6 months, progressing CKD was noted in 33 of the remaining 53 patients, confirming a significant burden of CKD following severe AKI. Serum creatinine and cystatin C emerged as the primary predictors in the feature importance analysis for CKD [24]. Furthermore, augmenting the training set with unlabelled data improved mortality predictions, suggesting the potential utility of partially labelled datasets.

Using EHR data derived from 350,345 cases across three hospitals, the final AKI recovery analysis included 7,041 patients in the internal cohort and 2,929 patients in the external cohort. A Categorical Boosting (CatBoost) machine learning model was trained to predict recovery of renal function, defined as a 33% decrease in serum creatinine from AKI onset or return to baseline. The model achieved internal and external AUROCs of 0.786 and 0.732, respectively, using data prior to 2023. The model also extended AI-based prediction beyond the intensive care setting to general ward patients, a group historically underrepresented in previous AKI prediction studies [2].

4.4. Deep Learning: Recurrent Neural Networks and Long Short-Term Memory

The time-series nature of sequential clinical data provides a unique benefit to deep learning architectures, especially recurrent neural networks (RNNs) with long short-term memory (LSTM) units, which can leverage temporal dependencies. In contrast to traditional ML algorithms, which consider features from only a single time point in their learning process, RNN-LSTM models utilise all collected clinical variables throughout the trajectory—something that is natural for modelling the dynamic progression of kidney injury and recovery. In a study of sepsis-associated AKI, RNN-LSTM, decision tree, and logistic regression models were evaluated for prediction of AKD, defined according to ADQI-16 criteria as kidney dysfunction persisting beyond 7 days and up to 90 days after AKI. In the development cohort of 209 patients, AUROCs were 1.000 for RNN-LSTM, 0.954 for decision trees, and 0.728 for logistic regression. In the external MIMIC-III validation cohort of 509 patients, the corresponding AUROCs were 1.000, 0.872, and 0.717, respectively [25]. The change in non-renal SOFA score from day 1 to day 3 was an important predictor of AKD [25]. Because this study focused on sepsis-associated AKI, its predictors and model performance should not be extrapolated directly to ischaemic, nephrotoxic, perioperative, or other AKI etiologies without external validation. These results suggest that machine learning strategies integrating longitudinal clinical information may enhance the prediction of AKD after sepsis-associated kidney disease. Several recent reviews have highlighted the broader potential of artificial intelligence and deep learning for predicting AKI, with AUROC values reported for outcomes such as end-stage renal disease and therapeutic response ranging from 0.85 to 0.96 [26]. Nonetheless, key issues around interpretability, fairness, and transparency continue to constrain clinical entrustment of these models [27].

4.5. Ensemble Learning Algorithms and Neural Network Architectures

Ensemble approaches have been evaluated for complex AKI outcomes. In a multicenter study using clinical data of 3 hospitals (2008–2020) that predicted dialysis dependence at 90 days after the first hospitalized dialysis (acute kidney injury-related dialysis [AKI-D]), an ensemble of 25 feedforward neural network models yielded an AUROC = 0.846, while an ensemble comprising 15 neural networks and 10 gradient-boosted decision trees yielded an AUROC = 0.865 for mortality at day 90 [28]. Baseline creatinine (obtained ≥90 days prior to dialysis) was the strongest predictor of dialysis dependence, whereas pre-dialysis creatinine was most predictive of survival [28]. In the cited multicentre AKI-D cohort, the ensemble models showed higher discrimination than logistic regression within that dataset; this finding supports further evaluation of model aggregation but should not be interpreted as universal superiority.

A multicenter machine learning model for AKI management in the general ward was developed using data from three hospitals in South Korea with a total of 135,068 patients, receiving AUROC scores of 0.905 and

0.886 for predicting AKI in internal and external cohorts, respectively, while achieving AUROCs of 0.820 and 0.783 for predicting AKD (here defined as non-recovery within seven days) [3]. These findings were strengthened by the use of a refined operational AKI definition, which mitigated misclassification caused by transient creatinine fluctuations. An explainable tree-based machine learning model was developed and externally validated to predict acute kidney injury stage 3 persistence (≥72 hours) in multinational data, achieving AUROCs of 0.94 in the US and 0.85 in the Italian cohorts among 7,759 ICU patients [16]. This multinational study demonstrates the potential value of cross-national validation of AI models for AKI prediction.

4.6. Prediction of Survival-Free from Dialysis and Renal Recovery

An important clinical challenge is to predict the recovery of renal function after AKI, as the status of recovery is fundamental for prognosis in patients with impaired kidneys. An interpretable XGBoost model to predict dialysis liberation before hospital discharge was created in 1,381 ICU patients with AKI requiring dialysis, among whom renal recovery occurred in 27.3% [29]. The AUROC was 0.85, and the area under the precision–recall curve was 0.69. Important predictors included urine output, the Charlson comorbidity index, trends in vitals (including respiratory rate and SpO_2), and lactate levels [29]. By considering how physiological variables change over time, these findings demonstrate the potential utility of dynamic clinical data for improving prediction.

Multiple machine learning algorithms were tested to predict 30-day mortality in severe AKI patients undergoing prolonged intermittent renal replacement therapy (PIRRT) based on data from the first 493 patients enrolled, with a 30-day mortality of 51.93% and a kidney recovery rate of 30.43% [30]. Within this cohort, different algorithms achieved the highest observed performance for different endpoints: Naïve Bayes for survival, Random Forest for 30-day renal recovery, and K-Nearest Neighbour for 90-day renal recovery [30]. The principal predictors were serum sodium, total protein, lactate dehydrogenase, phosphorus and thrombin time, among other clinical variables, highlighting the role of readily available biochemical parameters for prognostic stratification [30].

A meta-analytic synthesis of 27 studies showed higher accuracy for mortality than kidney function recovery prediction with machine learning models, and pooled analysis [31] indicated that traditional approaches were significantly outperformed for mortality by machine learning (AUROC 0.831 versus 0.772), but the study did not focus specifically on AKI survivor populations. The most predictive factors for recovery were age, serum creatinine, AKI stage, estimated glomerular filtration rate, and comorbidities [31]. If these results are confirmed in other studies, it would suggest that renal recovery in patients with AKI can be harder to predict than mortality.

4.7. Etiology- and Population-Specific AI Prediction After AKI

Available evidence includes distinct AKI contexts, including COVID-19-associated AKI, sepsis-associated AKI, severe AKI requiring renal replacement therapy, older adults, and paediatric populations. Model performance and predictor importance from one clinical phenotype should not be assumed transportable to another without external validation. Four machine learning algorithms (SVM, random forest, logistic regression, and gradient boosting) were tested to predict CKD progression among 131 discharged COVID-19 patients with AKI, achieving AUC-ROC values of 0.826 (SVM), 0.828 (random forest), 0.840 (logistic regression), and 0.841 (gradient boosting) [32]. Given the modest single-centre cohort and internal validation design, these findings require independent external validation [32]. Machine learning–based AKD and mortality prediction has also been used in older patients due to the increased susceptibility of geriatric populations to complications following AKI [33]. The paediatric applications of artificial intelligence have been comprehensively reviewed and included XGBoost, logistic regression, support vector machines, decision trees, as well as naïve Bayes, which were utilised in addition to risk stratification tools, including the Renal Angina Index and NINJA programme, demonstrating utility in AKI prediction and downstream consequences [34].

5. MODEL INTERPRETABILITY, POST-HOC EXPLAINABILITY, AND BLACK-BOX LIMITATIONS

Clinical adoption of AI prediction models critically depends on their interpretability—defined as the ability to offer transparent, clinically meaningful explanations for individual predictions. Based on cooperative game theory, the SHapley Additive exPlanations (SHAP) framework has been applied in AKI prediction studies as a common approach to estimating global feature importance rankings and patient-level explainability [2, 10, 20-22]. Some studies have reported that LIME and partial dependence plots are complementary tools for visualising predictions at the individual level [23, 29].

Interpretability differs substantially across modelling approaches. Conventional regression models and simple decision trees can provide comparatively direct structural relationships between predictors and model output, whereas random forests and gradient-boosted ensembles are less transparent because predictions arise from many interacting trees. Deep neural architectures, including RNN-LSTM models, are inherently more opaque because learned representations and temporal interactions are distributed across numerous parameters. In critical-care settings, this opacity can complicate error auditing, identification of spurious associations, recognition of dataset shift, and assessment of subgroup-specific failure, thereby increasing the risk of inappropriate reliance or automation bias.

Several variables recur across models predicting different post-AKI renal and clinical outcomes. However, feature importance is model-, endpoint-, population-, and

dataset-dependent and should not be interpreted as a universal causal hierarchy for AKI-to-CKD progression. Serum creatinine, both its absolute values and its early temporal dynamics (*i.e.*, changes from day 1 to day 3), has been widely reported as an important predictor across multiple models [21, 22, 28]. In a similar vein, blood urea nitrogen and its trajectory during the early post-AKI period were commonly identified as important predictors of outcomes [20-22]. Some variables, such as age, were consistently found to be important predictors across other studies and to reflect the association between age, comorbidity burden, and adverse outcomes [20, 24, 31]. Assessment of vasopressor use and mechanical ventilation as key predictors of MAKE30 and mortality mirrors the severity of illness and haemodynamic instability [10, 20]. Urine output and corresponding time trends were also strong predictors, especially in models of dialysis liberation and recovery [28, 30]. Additional predictive value, particularly in patients with severe AKI who required renal replacement therapy [20, 30], was demonstrated by indices of comorbidity (Charlson Comorbidity Index), as well as serum albumin and lactate parameters, and electrolyte disturbances (*i.e.*, sodium and phosphorus). The principal characteristics, predictive performance, validation strategies, interpretability approaches, and key predictors of the major AI/ML studies reviewed are summarised in Table **1**.

Post-hoc tools such as SHAP and LIME can help describe feature attribution and local model behaviour for a particular dataset or patient, but they do not make a black-box model inherently interpretable and should not be treated as proof of causal or biological mechanisms. Their explanations may vary with correlated predictors, background/reference data, model specification, and changes in the underlying patient population; therefore, they do not by themselves guarantee reliability, causal validity, transportability, or clinical trustworthiness. Accordingly, post-hoc explainability should complement rather than replace calibration assessment, external validation, subgroup performance assessment, clinical reasoning, prospective evaluation, and, where feasible, the use of inherently interpretable models in high-stakes clinical decisions [27, 35-37].

Table 1. Summary of Key AI/ML studies predicting renal and clinical outcomes across the AKI–AKD–CKD continuum.

Ref.	Design / Population / AKI Context	ML Algorithm(s)	Prediction Endpoint	AUROC / Performance	Sample Size	Validation Type	Interpretability	Key Predictors
[24]	Prospective; ICU; AKI stage 3	RF; RSF; Survival XGBoost	CKD at 3 and 6 months; mortality	AUPR: 0.895 / 0.848; C-index: 0.825	101 (+123)	Internal 5-fold CV; external	MDI	SCr, cystatin C, age, APACHE-II
[3]	Retrospective multicenter; general-ward AKI	ML framework (multiple)	AKI (3-day); AKD (7-day non-recovery)	AKI: 0.905 / 0.886; AKD: 0.820 / 0.783	135,068	External (3 hospitals)	Feature analysis	Refined SCr, vital signs, laboratory trends
[25]	Retrospective; sepsis-associated AKI, ICU	RNN-LSTM; decision trees; LR	AKD (ADQI-16 criteria)	LSTM: 1.00; DT: 0.954; LR: 0.728	209 + 509	External (MIMIC-III)	Feature importance	Δ non-renal SOFA (day 1–day 3)
[10]	Multicenter; ICU incident AKI	LR; RF; SVM; XGBoost	Mortality; MAKE (120-day)	Mortality: 0.79; MAKE: 0.78	9,587	10-fold CV + external	SHAP	71→15/14 features; SOFA/KDIGO comparators
[21]	Retrospective; elderly ICU AKI	LightGBM	AKD risk; in-hospital mortality	AKD: 0.844–0.853; mortality: 0.861–0.868	3,542 / 2,661	Bootstrap + external	SHAP + PDP	Cr day 3, sepsis, ΔBUN, DBP, HR
[20]	Retrospective; older hospitalized AKI	XGBoost + Boruta	MAKE30	Training: 0.868; internal: 0.823; external: 0.744	4,266 + 11,864	Internal + external (MIMIC-IV)	SHAP	Vasopressors, MV, BUN, RDW, albumin
[29]	Retrospective; ICU AKI-D	XGBoost	Dialysis liberation	AUROC: 0.85; AUPRC: 0.69	1,381	Temporal test set	SHAP + LIME	Urine volume, CCI, vital-sign derivatives, lactate
[19]	Retrospective; paediatric AKI	XGBoost *vs*. LR	MAKE30; 90-day adverse outcomes	MAKE30: 0.810; 90-day: 0.851	1,394	Internal split-set	SHAP; web calculator	Age, SCr, comorbidities
[2]	Multicenter; general-ward AKI	CatBoost	Renal function recovery	Internal: 0.786; external: 0.732	7,041 + 2,929	External (3rd hospital)	SHAP	43 features; SCr dynamics, comorbidities
[28]	Multi-hospital; AKI-D	Ensemble 25 FFN + 10 GBT	90-day dialysis dependence; mortality	Dialysis: 0.846; survival: 0.865	1,080 / 2,358	Multicenter (3 hospitals)	Feature importance	Baseline Cr, pre-dialysis Cr
[16]	Multinational; ICU AKI stage 2–3	Explainable tree regressor	Persistent AKI stage 3 (≥72 h)	US external: 0.94; Italian external: 0.85	7,759	Multinational external	Explainable model	Real-time clinical variables
[32]	COVID-19-associated AKI	SVM; RF; LR; boosting	CKD progression	SVM: 0.826; boosting: 0.841	131	Internal CV	Variable selection	COVID-specific clinical post-AKI variables

(Table 1) cont.....

Ref.	Design / Population / AKI Context	ML Algorithm(s)	Prediction Endpoint	AUROC / Performance	Sample Size	Validation Type	Interpretability	Key Predictors
[30]	Retrospective; severe AKI on PIRRT	NB; RF; KNN (by endpoint)	Survival; 30- and 90-day renal recovery	Variable by endpoint	493	Internal	Correlation analysis	Na, total protein, LDH, phosphorus

Abbreviations: AUROC, Area Under Receiver Operating Characteristic; AUPR, Area Under Precision-Recall Curve; AKD, Acute Kidney Disease; AKI-D, AKI requiring Dialysis; MAKE30, Major Adverse Kidney Events within 30 days; CCI, Charlson Comorbidity Index; Cr, Creatinine; BUN, Blood Urea Nitrogen; MV, Mechanical Ventilation; FFN, Feedforward Network; GBT, Gradient Boosted Trees; SCr, Serum Creatinine; CysC, Cystatin C; RSF, Random Survival Forest; LR, Logistic Regression; PDP, Partial Dependence Plot; MDI, Mean Decrease in Impurity; NB, Naive Bayes; KNN, K-Nearest Neighbor; RDW, Red Blood Cell Distribution Width; DBP, Diastolic Blood Pressure; HR, Heart Rate; LDH, Lactate Dehydrogenase; PIRRT, Prolonged Intermittent Renal Replacement Therapy; CV, Cross-Validation; Na, Sodium.

6. EXTERNAL VALIDATION, GENERALIZABILITY AND METHODOLOGICAL QUALITY

The clinical utility of any AI prediction model ultimately rests on its generalizability across populations and settings beyond the development cohort. In a systematic review and meta-analysis of machine learning studies for AKI prediction (n = 46), risk of bias was assessed using the TRIPOD, CHARMS, and PROBAST guidelines, which identified 39 studies (≈85%) as being at high risk of bias with models exhibiting limited interpretability and reproducibility that were externally validated in only eight studies (≈17%; five simultaneously conducted model development and validation; three had been validated) [17]. A later systematic review of 47 ICU-based studies found that only 14 (~30%) underwent external validation, and only 3 were considered to have a low risk of bias [18].

More recent methodological guidance has specifically addressed prediction models developed with regression and machine-learning methods. TRIPOD+AI provides updated reporting recommendations for the transparent description of data sources, predictors, outcomes, model development, performance, validation, and clinical applicability, and supersedes the original TRIPOD 2015 checklist for such prediction-model studies. PROBAST+AI updates risk-of-bias, quality, and applicability assessment for prediction models using regression or AI methods, with domains addressing participants and data sources, predictors, outcomes, and analysis. These frameworks are particularly relevant when appraising AI models whose apparent performance may be influenced by data provenance, preprocessing, model tuning, calibration, and external validation [38, 39].

Cross-institutional generalizability has come under explicit scrutiny. The Generalised DeepMind AKI prediction model trained on a predominantly male US Veterans population showed substantial sex-specific performance differences when validated in other cohorts, with the model performing worse for females due to miscalibration and lower discrimination [40]. This finding highlights the need for systematic fairness assessment across demographic subgroups—information that is largely absent in most published AI-AKI studies. Contemporary methodological guidance includes PROBAST+AI for quality, risk-of-bias, and applicability assessment and TRIPOD+AI for transparent reporting of prediction-model studies using regression or machine-learning methods. CONSORT-AI and SPIRIT-AI remain complementary guidance for interventional AI trials and protocols. TRIPOD-LLM provides additional reporting guidance specifically for studies developing, tuning, prompting, or evaluating large language models; it is not directly applicable to the non-LLM prediction models reviewed here [41]. The distance between model development and clinical deployment remains substantial: externally validated prediction models intended for real-world use are extremely rare [42]. Direct comparison of algorithm performance across studies is therefore problematic. An AUROC for XGBoost in one centre should not be ranked against an AUROC for random forest, LSTM, or another method in a different cohort when case mix, outcome prevalence, feature availability, missing-data handling, preprocessing, and validation differ. Even when the same algorithm is used, models fitted to different databases can learn different parameters and decision structures. Claims of algorithmic superiority should therefore be restricted to head-to-head comparisons within the same dataset or to replicated external-validation evidence.

7. INTEGRATION OF NOVEL BIOMARKERS WITH AI APPROACHES

Novel biomarkers and AI for AKI prediction have been evolving in parallel, with both approaches being investigated for improving AKI prediction [14]. In the RUBY study, urinary CCL14 emerged as the best biomarker of persistent stage 3 AKI (AUROC 0.83), significantly outperforming urinary KIM-1, plasma cystatin C, and urinary NGAL [11]. This biomarker has been linked to inflammatory pathways involved in kidney injury and repair. A critical comparison of ML and biomarkers was undertaken using an explainable ML model and urinary CCL14 for predicting persistent sepsis-associated AKI in retrospective or prospective cohorts [16]. The ML model used routinely available clinical data and outperformed or matched the analytical performance of urinary CCL14 [16]. Future hybrid strategies that combine biomarker information with ML may improve prediction of specific post-AKI renal outcomes, but their incremental value requires external validation.

8. CLINICAL IMPLEMENTATION: DECISION SUPPORT AND TRANSLATIONAL CHALLENGES

Ultimately, this area of research aims to translate AI prediction models into clinically deployed tools that enhance patient outcomes. AI applications in kidney care span several tasks, from prediction of incident AKI to post-

AKI risk stratification. Because incident AKI prediction is not equivalent to prediction of AKD or CKD progression, evidence from early-warning models should be interpreted separately from models of persistent kidney dysfunction, recovery, or long-term CKD [3]. Recent studies have confirmed the benefits of electronic AKI alerts and decision-support tools that monitor nephrotoxin exposures [13, 14] or provide kidney-focused care bundles to patients.

Progress has also been made toward making these tools clinically available. Several web applications for AKD prediction in elderly patients [21], paediatric AKI outcomes [19], and AKI/AKD prediction in gout patients [23] have been developed and deployed to bridge the gap between model development and bedside application. However, multiple barriers remain. The RE-AIM (Reach, Efficacy, Adoption, Implementation and Maintenance) framework has been suggested as a means of systematically assessing translational potential [43]. The broader question of whether ML in kidney disease is a genuine clinical advance or more a matter of premature enthusiasm remains an active discussion; the dearth of randomised evidence demonstrating improved patient-centred outcomes represents the largest evidentiary gap [44-46].

9. EMERGING DIRECTIONS AND FUTURE PERSPECTIVES

However, several emerging approaches may further advance AI-based prognostication across the AKI-AKD-CKD continuum and improve transportability across institutions. Federated learning (FL) enables institutions to collaboratively train models while preserving patient data, with only model parameters shared between sites [47]. One advantage of this approach is that it may alleviate the limitations of single-centre studies and help build more generalizable models. Complementary approaches include privacy-preserving solutions such as homomorphic encryption and blockchain technology [37].

A second important frontier is dynamic temporal modelling. There is evidence that temporal patterns in clinical data provide predictive information beyond static snapshots, as demonstrated in AKD prediction using RNN-LSTM models [25], prediction of dialysis liberation using vital sign trajectories [29], and prognostic assessment using delta creatinine and delta BUN measurements [21]. Future models incorporating continuous physiologic monitoring and real-time laboratory data may enable adaptive risk assessment during post-AKI recovery or progression.

Given its complexity, subphenotyping of AKI has been utilised to establish patient clusters with statistically disparate pathophysiological features and outcomes, thereby aiding precision nephrology approaches [27]. AI-based risk prediction could also change post-discharge surveillance for AKI survivors, many of whom currently have either no nephrology follow-up or are managed in outpatient settings after CKD has been established, despite their high future risk [5]. The inclusion of race-free CKD-EPI equations in predictive models may further improve the equity of long-term risk prediction [48].

10. STUDY LIMITATIONS

Although significant progress has been achieved to date, several important limitations currently hamper the clinical utility of AI-driven prognostic models across the AKI-AKI-CKD continuum. The majority of included studies were retrospective cohort designs, which are highly susceptible to selection and survivor bias. A heavy reliance on US-centric and high-income country datasets compromises global generalizability [1, 17]. The lack of standardised outcome definitions, prediction windows, and performance benchmarks limits comparability between studies [17, 18]. Additionally, the included studies span heterogeneous AKI etiologies and clinical settings and evaluate non-equivalent endpoints, including incident AKI, persistent AKI/AKD, renal recovery, dialysis dependence, MAKE, mortality, and CKD. Consequently, narrative cross-study comparisons should not be interpreted as head-to-head algorithm comparisons or as evidence that all included outcomes represent CKD transition. Calibration evaluation is often neglected in favour of discrimination measures alone [17]. Evaluation of fairness across demographic subgroups is still in its infancy [40]. Real-world problems with EHR data quality, missingness, and cross-system harmonisation remain major challenges. Incorporating AI tools into clinical workflows requires navigating alert fatigue, clinician trust, regulatory oversight, and liability [35, 45]. Most importantly, there is a lack of high-quality randomised evidence demonstrating improvements in patient-centred outcomes.

CONCLUSION

In summary, AI and ML show promise for prognostication across the AKI–AKD–CKD continuum, but the current evidence comprises heterogeneous prediction tasks rather than a single modelled transition from AKI to CKD. Studies have evaluated persistent AKI/AKD, renal recovery, dialysis dependence, MAKE, mortality, and CKD at 3–6 months using gradient boosting, random forests, recurrent neural networks, and ensemble methods. Although several studies report higher discrimination than conventional comparators within their own datasets, no algorithm can presently be considered universally superior because performance varies with AKI etiology, population, endpoint, feature set, and validation strategy. Serum creatinine dynamics, blood urea nitrogen, age, vasopressor use, urine output, and comorbidity burden were recurrent predictors across several studies, although their importance varied by population, endpoint, and model. SHAP, LIME, and other post-hoc tools can aid model explanation but should not be considered substitutes for inherent interpretability, external validation, calibration, and prospective clinical evaluation.

Long-term CKD after AKI remains an important clinical outcome, but it should be distinguished from earlier AKD, mortality, dialysis-related outcomes, and renal recovery. Future work should therefore use clearly defined and clinically meaningful endpoints, stratify or account for AKI

etiology where possible, and prioritise prospective multicentre external validation, calibration, fairness assessment, and evaluation of clinical utility.

However, substantial gaps remain between model development and clinical implementation. Essential priorities are prospective multicentre validation, dynamic real-time modelling, biomarker–AI integration, federated learning for privacy-preserving collaboration, comprehensive equity assessment, and, most importantly, randomised controlled trials that demonstrate whether AI predictions translate into better patient outcomes. Given rigorous prospective validation, AI-based prediction tools have the potential to improve the care of AKI survivors and reduce the long-term burden of chronic kidney disease.

AUTHORS' CONTRIBUTIONS

The authors confirm their contributions to the paper as follows:

M.A: Conceptualization, literature review, data curation, writing – original draft; V.N.: Supervision, validation, critical review, editing. Both authors reviewed the results and approved the final version of the manuscript.

LIST OF ABBREVIATIONS

ADQI	=	Acute Disease Quality Initiative
ADQI-16	=	Acute Disease Quality Initiative 16
AI	=	Artificial Intelligence
AKD	=	Acute Kidney Disease
AKI	=	Acute Kidney Injury
AKI-D	=	Acute Kidney Injury Requiring Dialysis
APACHE-II	=	Acute Physiology and Chronic Health Evaluation II
AUC-ROC	=	Area Under the Receiver Operating Characteristic Curve
AUROC	=	Area Under the Receiver Operating Characteristic Curve
AUPR	=	Area Under the Precision-Recall Curve
AUPRC	=	Area Under the Precision-Recall Curve
BUN	=	Blood Urea Nitrogen
CCI	=	Charlson Comorbidity Index
CCL14	=	C-C Motif Chemokine Ligand 14
CHARMS	=	Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies
CKD	=	Chronic Kidney Disease
CKD-EPI	=	Chronic Kidney Disease Epidemiology Collaboration
CONSORT-AI	=	Consolidated Standards of Reporting Trials–Artificial Intelligence
COVID-19	=	Coronavirus Disease 2019
Cr	=	Creatinine
CysC	=	Cystatin C
CV	=	Cross-Validation
DBP	=	Diastolic Blood Pressure
DNA	=	Deoxyribonucleic Acid
DT	=	Decision Tree
EHR	=	Electronic Health Record
eICU	=	eICU Collaborative Research Database
ESRD	=	End-Stage Renal Disease
FFN	=	Feedforward Network
FL	=	Federated Learning
GBT	=	Gradient-Boosted Trees
GFR	=	Glomerular Filtration Rate
HR	=	Heart Rate
ICU	=	Intensive Care Unit
IGFBP7	=	Insulin-Like Growth Factor-Binding Protein 7
KIM-1	=	Kidney Injury Molecule-1
KDIGO – Kidney Disease	=	Improving Global Outcomes
KNN	=	K-Nearest Neighbor
LDH	=	Lactate Dehydrogenase
LIME	=	Local Interpretable Model-Agnostic Explanations
LightGBM	=	Light Gradient Boosting Machine
LLM	=	Large Language Model
LR	=	Logistic Regression
LSTM	=	Long Short-Term Memory
MAKE	=	Major Adverse Kidney Events
MAKE30	=	Major Adverse Kidney Events Within 30 Days
MDI	=	Mean Decrease in Impurity
MIMIC-III	=	Medical Information Mart for Intensive Care III
MIMIC-IV	=	Medical Information Mart for Intensive Care IV
ML	=	Machine Learning
MV	=	Mechanical Ventilation
Na	=	Sodium
NB	=	Naïve Bayes

NGAL	=	Neutrophil Gelatinase-Associated Lipocalin
NINJA	=	Nephrotoxic Injury Negated by Just-in-Time Action
PDP	=	Partial Dependence Plot
PIRRT	=	Prolonged Intermittent Renal Replacement Therapy
PROBAST	=	Prediction model Risk Of Bias Assessment Tool
PROBAST+AI	=	Prediction model Risk Of Bias Assessment Tool for Artificial Intelligence
RDW	=	Red Blood Cell Distribution Width
RE-AIM	=	Reach, Effectiveness, Adoption, Implementation, and Maintenance
RF	=	Random Forest
RNN	=	Recurrent Neural Network
RNN-LSTM	=	Recurrent Neural Network–Long Short-Term Memory
RSF	=	Random Survival Forest
SASP	=	Senescence-Associated Secretory Phenotype
SCr	=	Serum Creatinine
SHAP	=	SHapley Additive exPlanations
SOFA	=	Sequential Organ Failure Assessment
SPIRIT-AI – Standard Protocol Items	=	Recommendations for Interventional Trials–Artificial Intelligence
SpO_2	=	Peripheral Oxygen Saturation
SVM	=	Support Vector Machine
TGF-β	=	Transforming Growth Factor Beta
TIMP-2	=	Tissue Inhibitor of Metalloproteinases-2
TRIPOD	=	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis
TRIPOD+AI	=	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis–Artificial Intelligence
TRIPOD-LLM	=	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis–Large Language Models
US	=	United States
XGBoost	=	Extreme Gradient Boosting

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors express their sincere gratitude to J.K.K. Nattraja College of Pharmacy, J.K.K. Nattraja Educational Institutions, Kumarapalayam, Namakkal, Tamil Nadu–638183, for providing access to relevant literature and resources during the preparation of this review. The authors also extend their sincere thanks to the faculty members of the Department of Pharmacy Practice for their valuable support and assistance throughout the preparation of this research.

REFERENCES

[1] Susantitaphong P, Cruz DN, Cerda J, *et al*. World incidence of AKI. Clin J Am Soc Nephrol 2013; 8(9): 1482-93. http://dx.doi.org/10.2215/CJN.00710113 PMID: 23744003

[2] Cho NJ, Jeong I, Kim Y, *et al*. A machine learning-based approach for predicting renal function recovery in general ward patients with acute kidney injury. Kidney Res Clin Pract 2024; 43(4): 538-47. http://dx.doi.org/10.23876/j.krcp.23.330 PMID: 38934029

[3] Cho NJ, Jeong I, Ahn SJ, *et al*. Machine learning to assist in managing acute kidney injury in general wards: Multicenter retrospective study. J Med Internet Res 2025; 27: e66568. http://dx.doi.org/10.2196/66568 PMID: 40101226

[4] Coca SG, Singanamala S, Parikh CR. Chronic kidney disease after acute kidney injury: A systematic review and meta-analysis. Kidney Int 2012; 81(5): 442-8. http://dx.doi.org/10.1038/ki.2011.379 PMID: 22113526

[5] Heung M, Chawla LS. Predicting progression to chronic kidney disease after recovery from acute kidney injury. Curr Opin Nephrol Hypertens 2012; 21(6): 628-34. http://dx.doi.org/10.1097/MNH.0b013e3283588f24 PMID: 23010757

[6] Ferenbach DA, Bonventre JV. Mechanisms of maladaptive repair after AKI leading to accelerated kidney ageing and CKD. Nat Rev Nephrol 2015; 11(5): 264-76. http://dx.doi.org/10.1038/nrneph.2015.3 PMID: 25643664

[7] Yang L, Besschetnova T Y, Brooks C R, Shah J V, Bonventre J V. Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. Nat Med 2010; 16: 535-43. http://dx.doi.org/10.1038/nm.2144 PMID: 20436483

[8] Guo C, Dong G, Liang X, Dong Z. Epigenetic regulation in AKI and kidney repair: Mechanisms and therapeutic implications. Nat Rev Nephrol 2019; 15(4): 220-39. http://dx.doi.org/10.1038/s41581-018-0103-6 PMID: 30651611

[9] Chawla LS, Bellomo R, Bihorac A, *et al*. Acute kidney disease and renal recovery: Consensus report of the Acute Disease Quality Initiative (ADQI) 16 workgroup. Nat Rev Nephrol 2017; 13(4): 241-57. http://dx.doi.org/10.1038/nrneph.2017.2 PMID: 28239173

[10] Neyra JA, Ortiz-Soriano V, Liu LJ, *et al*. Prediction of mortality and major adverse kidney events in critically ill patients with acute kidney injury. Am J Kidney Dis 2023; 81(1): 36-47. http://dx.doi.org/10.1053/j.ajkd.2022.06.004 PMID: 35868537

[11] Hoste E, Bihorac A, Al-Khafaji A, *et al*. Identification and validation of biomarkers of persistent acute kidney injury: The RUBY study. Intensive Care Med 2020; 46(5): 943-53. http://dx.doi.org/10.1007/s00134-019-05919-0 PMID: 32025755

[12] Ghazi L, Farhat K, Hoenig MP, Durant TJS, El-Khoury JM. Biomarkers *vs* machines: The race to predict acute kidney injury. Clin Chem 2024; 70(6): 805-19. http://dx.doi.org/10.1093/clinchem/hvad217 PMID: 38299927

[13] Mistry NS, Koyner JL. Artificial intelligence in acute kidney injury: From static to dynamic models. Adv Chronic Kidney Dis 2021; 28(1): 74-82. http://dx.doi.org/10.1053/j.ackd.2021.03.002 PMID: 34389139

[14] Bajaj T, Koyner JL. Artificial intelligence in acute kidney injury prediction. Adv Chronic Kidney Dis 2022; 29(5): 450-60. http://dx.doi.org/10.1053/j.ackd.2022.07.009 PMID: 36253028

[15] Kellum JA, Romagnani P, Ashuntantang G, Ronco C, Zarbock A, Anders HJ. Acute kidney injury. Nat Rev Dis Primers 2021; 7(1): 52. http://dx.doi.org/10.1038/s41572-021-00284-z PMID: 34267223

[16] Zappalà S, Alfieri F, Ancona A, *et al*. Development and external validation of a machine learning model for the prediction of persistent acute kidney injury stage 3 in multi-centric, multi-national intensive care cohorts. Crit Care 2024; 28(1): 189. http://dx.doi.org/10.1186/s13054-024-04954-8 PMID: 38834995

[17] Vagliano I, Chesnaye NC, Leopold JH, Jager KJ, Abu-Hanna A, Schut MC. Machine learning models for predicting acute kidney injury: A systematic review and critical appraisal. Clin Kidney J 2022; 15(12): 2266-80. http://dx.doi.org/10.1093/ckj/sfac181 PMID: 36381375

[18] Shi T, Lin Y, Zhao H, Kong G. Artificial intelligence models for predicting acute kidney injury in the intensive care unit: A systematic review of modeling methods, data utilization, and clinical applicability. JAMIA Open 2025; 8(4): ooaf065. http://dx.doi.org/10.1093/jamiaopen/ooaf065 PMID: 40620479

[19] Deng YH, Luo XQ, Yan P, Zhang NY, Liu Y, Duan SB. Outcome prediction for acute kidney injury among hospitalized children *via* eXtreme Gradient Boosting algorithm. Sci Rep 2022; 12(1): 8956. http://dx.doi.org/10.1038/s41598-022-13152-x PMID: 35624143

[20] Luo XQ, Zhang NY, Deng YH, Wang HS, Kang YX, Duan SB. Major adverse kidney events in hospitalized older patients with acute kidney injury: Machine learning–based model development and validation study. J Med Internet Res 2025; 27: e52786. http://dx.doi.org/10.2196/52786 PMID: 39752664

[21] Li M, Han S, Liang F, *et al*. Machine learning for predicting risk and prognosis of acute kidney disease in critically ill elderly patients during hospitalization: Internet-based and interpretable model study. J Med Internet Res 2024; 26: e51354. http://dx.doi.org/10.2196/51354 PMID: 38691403

[22] Qin H, Wu Q, Tan J, *et al*. Development of a machine learning-based prediction model for acute kidney injury associated with respiratory failure in the intensive care unit. Clin Exp Med 2025; 25(1): 326. http://dx.doi.org/10.1007/s10238-025-01873-y PMID: 41137985

[23] Jiang S, Xu L, Li C, *et al*. Development and validation of risk prediction models for acute kidney disease in gout patients: A retrospective study using machine learning. Eur J Med Res 2025; 30(1): 660. http://dx.doi.org/10.1186/s40001-025-02939-z PMID: 40702529

[24] 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 2023; 13(1): 9864. http://dx.doi.org/10.1038/s41598-023-36782-1 PMID: 37331979

[25] He J, Lin J, Duan M. Application of machine learning to predict acute kidney disease in patients with sepsis associated acute kidney injury. Front Med 2021; 8: 792974. http://dx.doi.org/10.3389/fmed.2021.792974 PMID: 34957162

[26] Yuan S, Guo L, Xu F. Artificial intelligence in nephrology: Predicting CKD progression and personalizing treatment. Int Urol Nephrol 2026; 58(7): 2615-45. http://dx.doi.org/10.1007/s11255-025-04878-4 PMID: 41205136

[27] Ozrazgat-Baslanti T, Loftus TJ, Ren Y, Ruppert MM, Bihorac A. Advances in artificial intelligence and deep learning systems in ICU-related acute kidney injury. Curr Opin Crit Care 2021; 27(6): 560-72. http://dx.doi.org/10.1097/MCC.0000000000000887 PMID: 34757993

[28] Wang TH, Kao CC, Chang TH. Ensemble machine learning for predicting 90-day outcomes and analyzing risk factors in acute kidney injury requiring dialysis. J Multidiscip Healthc 2024; 17: 1589-602. http://dx.doi.org/10.2147/JMDH.S448004 PMID: 38628614

[29] Wang TJ, Huang CT, Wu CL, *et al*. Predictive approach for liberation from acute dialysis in ICU patients using interpretable machine learning. Sci Rep 2024; 14(1): 13142. http://dx.doi.org/10.1038/s41598-024-63992-y PMID: 38849453

[30] Wei W, Cai Z, Chen L, Yuan W, Fan Y, Rong S. Short-term prognostic models for severe acute kidney injury patients receiving prolonged intermittent renal replacement therapy based on machine learning. BMC Med Inform Decis Mak 2023; 23(1): 133. http://dx.doi.org/10.1186/s12911-023-02231-2 PMID: 37488514

[31] Lin Y, Shi T, Kong G. Acute kidney injury prognosis prediction using machine learning methods: A systematic review. Kidney Med 2025; 7(1): 100936. http://dx.doi.org/10.1016/j.xkme.2024.100936 PMID: 39758155

[32] Gracida-Osorno C, Molina-Salinas GM, Góngora-Hernández R, Brito-Loeza C, Uc-Cachón AH, Paniagua-Sierra JR. Machine learning for predicting chronic renal disease progression in COVID-19 patients with acute renal injury: A feasibility study. Biomedicines 2024; 12(7): 1511. http://dx.doi.org/10.3390/biomedicines12071511 PMID: 39062084

[33] Wang X, Xu L, Guan C, *et al*. Machine learning-based risk prediction of acute kidney disease and hospital mortality in older patients. Front Med 2024; 11: 1407354. http://dx.doi.org/10.3389/fmed.2024.1407354 PMID: 39211338

[34] Raina R, Nada A, Shah R, *et al*. Artificial intelligence in early detection and prediction of pediatric/neonatal acute kidney injury: Current status and future directions. Pediatr Nephrol 2024; 39(8): 2309-24. http://dx.doi.org/10.1007/s00467-023-06191-7 PMID: 37889281

[35] Loftus TJ, Shickel B, Ozrazgat-Baslanti T, *et al*. Artificial intelligence-enabled decision support in nephrology. Nat Rev Nephrol 2022; 18(7): 452-65. http://dx.doi.org/10.1038/s41581-022-00562-3 PMID: 35459850

[36] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019; 1(5): 206-15. http://dx.doi.org/10.1038/s42256-019-0048-x PMID: 35603010

[37] Gogoi P, Valan JA. Machine learning approaches for predicting and diagnosing chronic kidney disease: Current trends, challenges, solutions, and future directions. Int Urol Nephrol 2025; 57(4): 1245-68. http://dx.doi.org/10.1007/s11255-024-04281-5 PMID: 39560857

[38] Collins GS, Moons KGM, Dhiman P, *et al*. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024; 385: e078378. http://dx.doi.org/10.1136/bmj-2023-078378 PMID: 38626948

[39] Moons KGM, Damen JAA, Kaul T, *et al*. PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ 2025; 388: e082505. http://dx.doi.org/10.1136/bmj-2024-082505 PMID: 40127903

[40] Cao J, Zhang X, Shahinian V, *et al*. Generalizability of an acute kidney injury prediction model across health systems. Nat Mach Intell 2022; 4(12): 1121-9. http://dx.doi.org/10.1038/s42256-022-00563-8 PMID: 38148789

[41] Gallifant J, Afshar M, Ameen S, *et al*. The TRIPOD-LLM reporting guideline for studies using large language models. Nat Med 2025; 31(1): 60-9.

http://dx.doi.org/10.1038/s41591-024-03425-5 PMID: 39779929

[42] Park S, Lee H. Acute kidney injury prediction models. Curr Opin Nephrol Hypertens 2019; 28(6): 552-9. http://dx.doi.org/10.1097/MNH.0000000000000536 PMID: 31356235

[43] Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. Am J Public Health 1999; 89(9): 1322-7. http://dx.doi.org/10.2105/AJPH.89.9.1322 PMID: 10474547

[44] Chan L, Vaid A, Nadkarni GN. Applications of machine learning methods in kidney disease. Curr Opin Nephrol Hypertens 2020; 29(3): 319-26. http://dx.doi.org/10.1097/MNH.0000000000000604 PMID: 32235273

[45] Yu X, Ji Y, Huang M, Feng Z. Machine learning for acute kidney injury: Changing the traditional disease prediction mode. Front Med 2023; 10: 1050255. http://dx.doi.org/10.3389/fmed.2023.1050255 PMID: 36817768

[46] Annalisa S, Marco F, Bibiana B, *et al*. Acute kidney injury prediction and prognostication using machine learning. Int Urol Nephrol 2026; 58(8): 3161-70. http://dx.doi.org/10.1007/s11255-025-04959-4 PMID: 41436719

[47] Rieke N, Hancox J, Li W, *et al*. The future of digital health with federated learning. NPJ Digit Med 2020; 3(1): 119. http://dx.doi.org/10.1038/s41746-020-00323-1 PMID: 33015372

[48] Inker LA, Eneanya ND, Coresh J, *et al*. New creatinine- and cystatin C–based equations to estimate GFR without race. N Engl J Med 2021; 385(19): 1737-49. http://dx.doi.org/10.1056/NEJMoa2102953 PMID: 34554658