REVIEW

# Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review

Greet De Vlieger[a], Kianoush Kashani[b,c], and Geert Meyfroidt[a]

**Purpose of review**

Acute kidney injury (AKI) frequently complicates hospital admission, especially in the ICU or after major surgery, and is associated with high morbidity and mortality. The risk of developing AKI depends on the presence of preexisting comorbidities and the cause of the current disease. Besides, many other parameters affect the kidney function, such as the state of other vital organs, the host response, and the initiated treatment. Advancements in the field of informatics have led to the opportunity to store and utilize the patient-related data to train and validate models to detect specific patterns and, as such, predict disease states or outcomes.

**Recent findings**

Machine-learning techniques have also been applied to predict AKI, as well as the patients' outcomes related to their AKI, such as mortality or the need for kidney replacement therapy. Several models have recently been developed, but only a few of them have been validated in external cohorts.

**Summary**

In this article, we provide an overview of the machine-learning prediction models for AKI and its outcomes in critically ill patients and individuals undergoing major surgery. We also discuss the pitfalls and the opportunities related to the implementation of these models in clinical practices.

**Keywords**

acute kidney injury, artificial intelligence, machine learning, prediction

## INTRODUCTION

Acute kidney injury (AKI) frequently occurs in hospitalized patients, especially in the ICU [1]. AKI has a strong association with increased short-term and long-term morbidity and mortality, and increased use of healthcare resources [2,3]. The consensus diagnosis and staging of AKI according to increased serum creatinine (SCr) or reduced urinary output, has undergone slight modifications from the original RIFLE and AKI Network criteria [4,5] to the most recent Kidney Disease Improving Global Outcome (KDIGO) criteria [6]. No specific treatment has proven to change the course of AKI, and management is mainly supportive of preventing further deterioration. Progressive decline in kidney function, fluid overload, or metabolic complications can be treated with renal replacement therapy (RRT) when the kidney cannot satisfy the demands of osmolar and fluid load [7]. On the contrary, an increase in SCr is a late and insensitive marker of the underlying decline in glomerular filtration rate (GFR), and significant damage has already occurred at the time of diagnosis [8]. Consequently, the window to adapt treatment to prevent AKI is short.

The development of AKI depends on many parameters, including patient characteristics such as age and comorbidities, the causal event, the host response and impact on vital functions, and the initiated treatment and resulted responses [9]. Recent advances in the field of informatics and the ability to collect and store unlimited data have led to the increased use of electronic health records in the advent of clinical decision supports. The tremendous computational capabilities of new technologies could complement and enhance human performance in overview and interpret all these data

[a]Clinical Division and Laboratory of Intensive Care Medicine, Academic Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium, [b]Division of Nephrology and Hypertension and [c]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA

Correspondence to Greet De Vlieger, Clinical Division and Laboratory of Intensive Care Medicine, Academic Department of Cellular and Molecular Medicine, KU Leuven, Herestraat 49, 3000 Leuven, Belgium.
Tel: +00 32 16344017; e-mail: Greet.devlieger@uzleuven.be

Curr Opin Crit Care 2020, 26:000–000

DOI:10.1097/MCC.0000000000000775

## KEY POINTS

- Artificial intelligence is a helpful tool to predict the occurrence of AKI in critically ill patients.
- Several models have been developed and are ready for external validation.
- Artificial intelligence-based predictions may be helpful to indicate in whom biomarkers for AKI may be more powerful.
- Further research is needed to assess the impact of AKI predicting models on the outcome.

on a continuous basis [10]. Several model building and machine learning techniques have been developed and are increasingly applied in ICU [11]. Models can be constructed by regression learning for continuous outcome parameters [11]. For binary outcomes, 'classification learning' is used to train models to classify patterns based upon characteristic variables (Table 1). The selection of independent variables can be made by Least Absolute Shrinkage and Selection Operator. Regression and classification techniques can both be applied in big datasets. The models' performance is evaluated by discrimination or area under the receiver-operating curve (AUROC), calibration, and area under the precision curve [11,12]. Here we provide an overview of artificial intelligence-based tools created to predict AKI in critically ill patients and in patients undergoing major surgical procedures and ICU outcomes in patients with AKI.

## METHODS

We searched PubMed using Medical Subject Headings terms 'artificial intelligence' or 'machine learning', and 'acute kidney injury'. All 51 reports were analyzed, and the reference list was searched for additional relevant reports. Also, we evaluated the citing articles of all related articles. Then, we classified the articles into the type of admission (ICU, perioperative or general ward) and the predicted outcome. We included only the reports with a detailed description of the model evaluation in ICU patients and patients undergoing major surgery (Tables 2 and 3).

## PREDICTION OF ACUTE KIDNEY INJURY

### ICU risk models

We found eleven AKI prediction models for ICU patients [13–16,17[*],18,19[*],20,21[**],22,23]. While most studies included mixed ICU populations [13–16,17[*],18,19[*],20], two included all hospitalized patients but reported the model performance of ICU patients separately [17[*],21[**]], and two studies focused on patients with severe burns [22,23]. Six studies used SCr to define AKI [13–16,17[*],21[**]], one also included estimated GFR [18], and four studies used both SCr and urinary output [13,19[*],22,23]. The baseline SCr was not determined uniformly. While some studies used the average SCr from 365 to 1 day before admission [13–16,21[**]], others used SCr upon admission [17[*]], back-calculated SCr based upon the modified diet and renal disease (MDRD) formula [18], or any measured SCr throughout the hospitalization [20]. Three studies did not report the definition of baseline SCr [19[*],22,23]. In articles that reported the rate of missing SCr baseline values, the percentage of back-calculated SCr by the MDRD formula was 21.8–30.0% [14,15]. The occurrence of AKI was predicted during the first 24–48 h after ICU admission in five studies [14,18,20,22,23], two studies estimated the risk every 15 min [13,19[*]], two studies made a daily prediction [17[*],21[**]], one study on four different time points early during ICU admission [15], and one study did not report it [16]. The prediction window was 1 week in five studies [14,15,21[**],22,23], while two studies predicted AKI in the following 72 h [16,18], one study within 48 h [17[*]], and three studies did not report it [13,19[*],20].

Three models were validated in external cohorts [14,15,23]. Malhotra *et al.* [14] developed the model locally and validated it in an external cohort with a higher incidence of AKI, which suggests that the populations were not entirely comparable. Despite differences in the incidence of AKI in two populations, the discrimination of the models was good in both development and validation cohorts. Flechet *et al.* [24] compared the model versus physicians' prediction in a prospective observational study in the same tertiary, teaching hospital, and included urinary output along with SCr to define AKI. The lower incidence of AKI in this cohort was related to excluding nonsurgical patients. They found a good discrimination and accuracy, for both the machine learning model and physicians' predictions. However, there was a delay in the physicians' risk determination as compared with the prediction time by the AKI predictor. The authors concluded, an automated alert is likely to predict AKI earlier as physicians are not able to process all parameters continuously. Moreover, the prediction from physicians with less experience demonstrated a lower discrimination and calibration. Lastly, Rashidi *et al.* [23] developed and validated an AKI prediction model among severely burned patients. However,

AQ10 Table 1. Overview of the most frequently applied machine learning techniques for acute kidney injury prediction

| Machine learning model | Description | Strengths | Caveats |
|---|---|---|---|
| NB | Link from target variable to nontarget variables<br>Assumes that all predictors are not dependent on each other | Easy and quick<br>Does not need large training datasets | Assumption that predictors are independent is very unlikely in most data<br>Notably poor in estimating probabilities |
| DT | Tree-shaped model that sequentially splits the data according to the most predictive attribute. Identification of a small set of variables that have high predictive power for the predicted parameter | Easily understandable<br>Robust to labeling errors and noise<br>Costs may be assigned to attributes | Do not always perform well<br>High risk of overfitting because of small random variations in the data |
| RF | Repeated number of DT, on slightly perturbed versions of the original dataset. Generates a combined prediction from the different DT | Fast<br>Can deal with missing data<br>Removing influence of small random variations<br>Better performance than DT | More difficult to interpret than a DT<br>High risk of overfitting because of small random variations in the data |
| GBT | A combined prediction from different DT. Typically, it will combine weakly performing DTs in an iterative way. As the boosting process continues, the new trees will focus more on the examples that were misclassified by previous trees | Reduces bias<br>Can convert a weak classifier into a stronger classifier | Difficult to interpret<br>Risk of overfitting can be overcome by limiting the allowed number of trees in the model |
| NN | Collection of processing units interconnected to increase the power over a single unit | Robust to errors, well suited for noisy examples<br>Frequently outperforms other ML techniques<br>Can be used for many types of data, including images, audio... | Risk of overfitting<br>Needs large datasets for training<br>Long training times<br>Difficult to interpret, 'black box' model |
| KNN | Classifies examples, and assigns them a value according to the plurality of their nearest neighbors. *K* refers to the (low) number of neighbors considered by the model | Robust to errors, well suited for noisy examples | Needs large datasets for training<br>Long training times<br>Determining the value of *K* is critical<br>High computational cost |
| SVM | Combination of several dimensions of binary classification | Outperform multivariate linear regression<br>Easy to train | Only binary classifiers, but solutions for regression exist |

AKI prediction is a classification learning tool. AKI, acute kidney injury; DT, decision trees; GBT, gradient boosted trees; KNN, *K*-nearest neighbors; ML, machine learning; NB, naïve Bayes; NN, neural networks; RF, random forest; SVM, support vector machines.

the sample size for both the development and validation of was small.

## Perioperative risk models

Major surgery is frequently complicated by AKI, which is associated with increased short-term and long-term mortality and morbidity [25,26]. During the latest decades, the awareness of perioperative AKI has increased, and guidelines for perioperative management have been published [27]. Although the KDIGO-based diagnosis of AKI is frequently made postoperatively in the intensive or postoperative care unit, in most cases, AKI develops already in the operating theatre.

Thottakkara *et al.* [28] were the first to report an machine learning prediction model for postoperative AKI. The input variables were preoperative information, without including perioperative hemodynamic parameters. The model was able to predict AKI in the first postoperative week with good discrimination. Bihorac *et al.* [29] further explored this model to develop and validate a score to predict mortality and major postoperative complications named MySurgeryRisk. The model has been compared with the prediction by clinicians in 150 patients and found to have significantly higher discrimination for predicting AKI than the clinicians' forecast [30]. Trainees misclassified patients more often as compared with attending physicians, but the differences were NS.

The same investigators added perioperative data to the model to build a dynamic machine learning algorithm [31]. Postoperative AKI before hospital

CE: Swati; MCC/260604: Total nos of Pages: 10;

MCC 260604

Table 2. Overview of the machine learning generated prediction models for acute kidney injury in ICU patients and in postoperative patients; for the need of renal replacement therapy; and for volume responsiveness and mortality in critically ill acute kidney injury patients

| | Study design | *n* | Region and time period | Patient population | Prediction | Baseline SCr | ML | Prediction frequency | Prediction window | Prediction incidence |
|---|---|---|---|---|---|---|---|---|---|---|
| Prediction of AKI in ICU | | | | | | | | | | |
| Ahmed 2015 [13] | Retrospective development | 482 | USA, Olmsted County, Mayo Clinic July 2010–December 2010 | Tertiary-mixed ICU | AKIN SCr and UO | Median of all values during 180 days prior to the admission or by calculation from the MDRD formula | MATLAB | Every 15 min | – | 30.0% |
| | Retrospective validation | 462 | USA, Rochester, Mayo Clinic January 2010–March 2010 | | | | | | 9.7 h | 40.0% |
| Malhotra 2017 [14] | Retrospective development | 573 | USA, UCSD, San Diego June 2006–December 2008 | ICU patients, first 48 h of admission | AKI KDIGO SCr | Mean of 7–365 days prior to admission, imputation when missing baseline (24%) | Multiple LR | Once, first 48 h of admission | 7 Days (median time 23.2 h) | 22.0% |
| | Retrospective internal validation | 144 | USA, UCSD, San Diego June 2006–December 2008 | | | Mean of 7–365 days prior to admission, imputation when missing baseline (24%) | | | 7 Days (median time 23.2 h) | 24.0% |
| | Prospective external validation | 1300 | USA, Olmsted County, Mayo Clinic January 2010–December 2010 | | | Mean of 7–365 days prior to admission, imputation when missing baseline (30%) | | | 7 Days (median time 24.4 h) | 45.0% |
| Flechet 2017 [15] | Retrospective development | 2123 | Belgium, Leuven August 2007–November 2010 | Tertiary-mixed ICU | AKI KDIGO SCr | Lowest in 3 months before admission 77.2%, MDRD formula in 21.8% | RF | Before and upon admission, after 1 day in ICU and after first 24 h in ICU | First week ICU stay | 27.7% AKI, 14.0% AKI 2–3 |
| | Retrospective validation | 2367 | Belgium, Leuven August 2007–November 2010 | | | Lowest in 3 months before admission in 77.1%, MDRD formula in 22.9% | | | | 29.2% AKI, 14.7% AKI 2–3 |
| Flechet 2019 [24] | Prospective validation | 252 | Belgium, Leuven 2018 | Tertiary surgical ICU | AKI KDIGO 2 or 3 (SCr and UO) | Lowest SCr in 3 months prior to and not including admission, when not available calculated from MDRD formula | | Upon admission, first morning in ICU and after 24 h | 27.1 h in admission cohort, 39.7 h in day 1 cohort and 39.7 h in day 1þ cohort | 12.0% |
| Mohamadlou 2018 [16] | Retrospective development and three-fold cross-validation | 48 582 | USA, Boston, MIMIC-III database 2001–2012 | ICU | AKI stage 2 or 3 NHS England AKI algorithm | Lowest value past week or median value from past 8 to 365 days | GBT | Not reported | Up till 72 h | 2.7% |
| Koyner 2018 [17&] | Retrospective development | 72 694 | USA, Chicago November 2008–January 2016 | Tertiary urban hospital, 28.9% of AKI in ICU | AKI KDIGO 2 SCr | Admission SCr | GBT | Daily | 48 h (median 41 h for AKI stage 2) | 14.4% AKI 3.5% AKI stage 2 |
| | Retrospective validation | 48 464 | | | | | | | | |
| Zimmerman 2019 [18] | Retrospective development | 23 950 | Israel, MIMIC III 2001–2012 | ICU patients without preexisting CKD or AKI | KDIGO SCr or GFR < 0.5 ml/kg/h for more than 6 h | Calculated MDRD formula | LR, RF, NN | On day 1 of ICU admission | Within 72 h of ICU admission | 16.5% |
| Chiofolo 2019 [19&] | Retrospective development | 4572 | USA, Minnesota October 2004–April 2011 | Mixed ICU | AKI AKIN SCr and UO | Not reported | RF | Every 15 min | | 30% (All stages) |
| | Retrospective validation | 1958 | | | | | | | | 30% (All stages) |
| Parreco 2019 [20] | Retrospective development 10-fold cross validation | 151 098 | USA 2014–2015 | ICU | KDIGO SCr | Daily serum creatinine in the ICU | GBT, LR, deep learning | Once, first 48 h of admission | | 5.6% |

Table 2 (*Continued*)

| | Study design | *n* | Region and time period | Patient population | Prediction | Baseline SCr | ML | Prediction frequency | Prediction window | Prediction incidence |
|---|---|---|---|---|---|---|---|---|---|---|
| Tomašev 2019 [21&&] | Retrospective training | 563 026 | USA, Veterans Affairs January 2011–September 2015 | Hospital admission | AKI KDIGO creatinine | Median of all values 1 year prior to admission | RNN | Continuously | 48 h | 13.4% in total cohort (not reported for ICU separately) |
| | Retrospective validation | 35 189 | | | | | | | | |
| | Retrospective calibration | 35 189 | | | | | | | | |
| | Retrospective test | 70 378 | | | | | | | | |
| Tran 2019 [22] | Retrospective development | 40 | Not reported | Adults with burns 2'20% TBSA | AKI KDIGO SCr and UO | Not reported | KNN | 1x within 24 h following burn ICU admission | First week of ICU stay, mean 42.7 h | 50.0% |
| | Retrospective validation | 10 | | | | | | | | |
| Rashidi 2020 [23] | Retrospective development | 50 | Not reported | Burns TBSA 2' 20% | AKI KDIGO SCr and UO | Not reported | LR, KNN, SVT, RF, NN | Once | First week of ICU stay | 50.0% |
| | Prospective validation | 51 | | Burns TBSA 2' 20% or nonburn trauma requiring surgery | AKI KDIGO SCr and UO | | | | First week of ICU stay, mean 71.5 h | 34.2% |
| Prediction of AKI after surgery | | | | | | | | | | |
| Thottakkara 2016 [28] | Retrospective development | 35 223 | USA, Florida January 2000–November 2010 | Major surgery | KDIGO SCr | Preoperative serum creatinine | LR, GAM, NB, SVT | Once preoperative | AKI in the first 7 days after surgery | 36.0% |
| | Retrospective validation | 15 095 | | | | | | | | |
| Bihorac 2019 [29] | Retrospective development | 41 166 | USA, Florida January 2000–November 2010 | Major surgery | KDIGO SCr | Preoperative serum creatinine | GAM, RF | Not reported | AKI during admission | 38.9% |
| | Retrospective validation | 10 291 | USA, Florida January 2000–November 2010 | | | | | Not reported | | |
| Brennan 2019 [30] | Retrospective validation | 150 | USA, Florida January 2000–November 2010 | Major surgery | KDIGO SCr | Preoperative serum creatinine | | Once preoperative | AKI in the first 7 days after surgery | 38.0% |
| Adhikari 2019 [31] | Retrospective development and five-fold cross-validation | 2038 | USA, Florida January 2000–November 2010 | Surgery | KDIGO SCr | Lowest value 7 days prior to admission, median creatinine 8–365 days prior to admission or calculated using the MDRD | RF, GAM | Perioperatively and immediately after surgery | 3 Days after surgery, 7 days after surgery, and overall hospitalization | 46.0% |
| | Retrospective validation | 873 | | | | | | | | |
| Lei 2019 [32] | Retrospective development | 25 616 | USA, Pennsylvania January 2014–April 2018 | Major noncardiac surgery | KDIGO SCr | Lowest creatinine value within 7 days before surgery, if not available the most recent value up to 365 days before surgery | Elastic net selection, GBT, RF | Prehospitalization, preoperative, perioperative | AKI 1 week after surgery | 10.4% |
| | Retrospective validation | 8505 | | | | | | | | 9.6% |
| | Retrospective test | 8494 | | | | | | | | 9.9% |

Table 2 (*Continued*)

| | Study design | *n* | Region and time period | Patient population | Prediction | Baseline SCr | ML | Prediction frequency | Prediction window | Prediction incidence |
|---|---|---|---|---|---|---|---|---|---|---|
| Lee 2018 [33] | Retrospective development | 848 | Korea, Seoul November 2004–December 2015 | Liver transplantation, adults | AKIN SCr | Most recent SCr before surgery | DT, RF, GBM, SVT, NB, multilayer perceptron and deep belief networks | At transplantation | First 2 postoperative days | 30.0% |
| | Retrospective validation | 363 | | | | | | | | 31.0% |
| Prediction of the need for RRT | | | | | | | | | | |
| Cronin 2015 [34] | Retrospective development and validation by bootstrapping | 1620 898 | USA, Veterans Affairs hospitals January 2003–December 2012 | Hospital admissions 48 h–30 days | RRT | Mean outpatient creatinine from 365 to 7 days before admission | LR, LASSO LR, RF | Not reported | 48 h–9 days after admission | 0.12% |
| Flechet 2017 [15] | Retrospective development | 2123 | Belgium, Leuven August 2007–November 2010 | Tertiary-mixed ICU | RRT | Lowest in 3 months before admission in 77.1%, MDRD formula in 22.9% | RF | Before and upon admission, after 1 day in ICU and after first 24 h in ICU | First week ICU stay | 7.3% |
| | Retrospective validation | 2367 | | | | | | | | 7.6% |
| Saly 2017 [35] | Retrospective development | 1098 | USA, Pennsylvania September 2013–April 2014 | AKI in hospital, 30.2% ICU | RRT | | RF | | 1 Week | 7.5% |
| | Retrospective validation | 1143 | | | | | | | | |
| Koyner 2018 [17 &] | Retrospective development | 72 694 | USA, Chicago November 2008–January 2016 | Tertiary urban hospital, 20.8% ICU admissions | RRT > 48 h after admission | NA | GBT | Daily | | 0.68% |
| | Retrospective validation | 48 464 | | | AKI KDIGO stage 2 creatinine | Admission SCr | | | | |
| Prediction of volume responsiveness in AKI patients in the ICU | | | | | | | | | | |
| Zhang 2019 [40] | Retrospective development | 5012 | USA, Boston June 2001–October 2012 | ICU patients with UO < 0.5 ml/kg/h for 6 h and fluid intake >5 l during next 6 h | Volume responsiveness (UO 2' 0.65 ml/kg/h next 12 h) | | LR and XGBoost with decision trees | Once | Next 18 h | 58.1% |
| | Retrospective validation | 1670 | | | | | | | | |
| Prediction of mortality in AKI patients in the ICU | | | | | | | | | | |
| Lin 2018 [41] | Retrospective development and five-fold cross-validation | 19 044 | USA, Boston 2001–2012 | ICU patients with AKI (KDIGO criteria) | In-hospital mortality | NA | RF, NN, SVM | Once | Hospitalization | 13.6% |

AKI, acute kidney injury; AKIN, AKI Network; CKD, chronic kidney disease; GAM, generalized additive model; GBT, gradient boosted trees; GFR, glomerular filtration rate; KDIGO, Kidney Disease Improving Global Outcome; KNN, *K*-nearest neighbors; LASSO, Least Absolute Shrinkage and Selection Operator; LR, logistic regression; MDRD, modified diet and renal disease; ML, machine learning; NN, neural networks; RF, random forest; RNN, recurrent neural networks; RRT, renal replacement therapy; SCr, serum creatinine; TBSA, total BSA; UO, urinary output.

discharge occurred in 46% of the patients, of whom 87% developed it within the first week after surgery. More than half of the patients were admitted to the ICU for 48 h or more. Adding the perioperative data improved the discrimination and accuracy of the prediction and resulted in a net reclassification index of 11.02% during the first 72 postoperative hours.

Lei *et al.* [32] developed an AKI prediction model for the first postoperative week in patients undergoing major, noncardiac surgery. They constructed three models using prehospitalization, preoperative, and perioperative variables. The model improvement was more significant when adding preoperative variables as compared with perioperative variables.

Finally, Lee *et al.* [33] developed a model to predict AKI within 48 h after liver transplantation. A model build with the gradient boosting technique had the best performance with an AUROC of 0.90.

## PREDICTION OF THE NEED FOR RENAL REPLACEMENT THERAPY

Four models predicted the need for RRT in hospitalized patients [15,17▪,34,35]. One study included only ICU patients [15], whereas in two others, 20 and 30% of the patients were admitted to the ICU [17▪,35], and one study included all hospitalized patients [34]. The time window of prediction varied from the next 48 h to the total duration of admission. Models discriminations were good, with AUROC between 0.82 and 0.96.

## OTHER OUTCOME PREDICTIONS IN ICU PATIENTS WITH ACUTE KIDNEY INJURY

### Volume responsiveness

Intravenous fluid challenges are often administered in critically ill patients to restore cardiac output and improve kidney function. On the other hand, fluid overload is more common in AKI patients and is associated with worse outcome [36,37]. Restrictive fluid management has shown to reduce the worsening of AKI in septic patients [38] but may increase the risk of developing AKI among patients undergoing major abdominal surgery [39]. Zhang *et al.* investigated whether artificial intelligence may predict fluid responsiveness in oliguric patients who received fluid resuscitation. Fluid responsiveness was defined as an increase in urinary output. In an internal validation, the model had an AUROC of 0.860 [40].

### Mortality

Lin *et al.* [41] developed a prediction model for mortality in ICU patients with AKI. Included comorbidities were limited to AIDS, metastatic cancer, and hematologic malignancy. As death is frequently secondary to comorbidities and the decision not to start RRT and/or withdraw treatment is inseparable from the patient's medical history, including more parameters, it may further improve the model.

## DISCUSSION

AKI often results from several exposures and presents more frequently in patients with comorbidities. Risk assessment for AKI also includes these current exposures, age, comorbidities, host response, treatments, and treatment response. Artificial intelligence can integrate all parameters and may be valuable for AKI prediction. Several models have been developed and validated to predict AKI in ICU and perioperative settings. Most of them show a good discrimination and accuracy in internal validation, and also three models that were validated in external cohorts show promising results [14,23,24].

Nevertheless, artificial intelligence can never replace physicians because machines are not able to integrate the predictions into a balanced clinical decision [42]. Instead, artificial intelligence must be considered as a rapid and efficient tool to detect patients at risk [43]. Several models can predict AKI in the next 48–72 h with good accuracy [15,16,17▪,18]. Models that predict the risk for a shorter period (in other words, closer window to AKI development) perform better but may lack the possibility to improve outcomes on short notice. Indeed, patients will potentially benefit from a correct prediction if action is taken early enough to allow effective preventive measures, so a considerable time window is needed to allow the physician to intervene and prevent further deterioration [44]. The question of whether machine learning models may improve outcomes must be evaluated in an 'impact study', that is a randomized controlled trial that compares a cohort that physicians have access to the model results versus a cord of standard of care [45,46].

Implementation of machine learning models in daily practice may help to improve the algorithm, as machine learning allows to improve the performance in the presence of additional curated information [42]. Therefore, continuous input of data is needed to recalibrate the model. As models likely overestimate the risk after improvement in medical care, the recalibration leads to continued accurate prediction [12,47].

Physicians must be aware of the pitfalls when a model is introduced into clinical practice. First, the quality of the prediction depends on the quality of

data (garbage in, garbage out) [48]. To apply artificial intelligence in clinical practice, data need to be available in real-time, curated, and assessed for accuracy and reliability, similar to their availability during development and validation. Second, initiation of the model will be most optimal if the model is used in a similar population to where it was built [49]. For instance, algorithms generated on databases of the (predominantly male) Veterans Affairs will initially misclassify women more frequently. This could be corrected by recalibration, provided data on female patients become available. This is also the reason why models must be validated in previously unseen populations, prospectively, and ideally in every hospital where the model is introduced. Besides, reports of prediction models must summarize the characteristics of the population in whom the model is built, which is very relevant for the physician to assess whether a model can be used in a particular setting. Recommendations for reporting prediction models are summarized in the Transparent Reporting of a multivariable prediction model for individual Prognosis or Diagnosis statement [12]. Third, the predicted outcome should be defined according to objective and measurable adjudicating criteria (e.g., length of stay, AKI [50], or death), and should be clinically relevant. A drop in performance may be expected if the predicted outcome is a medical intervention (e.g., RRT), especially when treatment strategies differ between hospitals. While the KDIGO criteria define AKI according to a rise in SCr or a drop in urinary output, most models only use the SCr criterion and some use nonstandardized definitions of baseline SCr. Preferably, the predicted outcome should have an established physiological and clinically relevant association with the predictors. For example, in a model for prediction of fluid responsiveness, an increase in urinary output was used to assess fluid responsiveness [40], even while such increase in urinary output is not generally accepted as a proper parameter of fluid responsiveness [51]. Artificial intelligence models are particularly attractive as they can also be used to predict other relevant clinical outcomes (e.g., ARDS, heart failure, liver failure, shock) or even the effect of specific treatments. Such predictions may propose potentially successful treatments to the clinician, which would increase the clinical applicability of machine learning models, according to a survey [52]. Fourth, ample research has been done on the warning threshold and alerting method for such AKI sniffers. The benefit of a high sensitivity must be out weighted to the risk of alarm fatigue caused by a low positive predictive value [53,54].

AQ7

AKI prediction models are often compared with laboratory biomarkers. It is well known that the performance of AKI prediction biomarkers varies in different populations [55,56], and may be more accurate in high-risk populations. Compared with biomarker tests, machine learning models could be used to predict AKI continuously without additional costs [13,19*]. This continuous risk evaluation may detect those high-risk patients, in whom further testing with biomarkers is indicated. Further research is needed to evaluate the performance of AKI biomarkers in high-risk patients identified by machine learning algorithms.

## CONCLUSION

Artificial intelligence is increasingly used in medicine. Large ICU databases are used to build machine learning prediction models for AKI and other outcomes. While several groups have developed models with acceptable to very good performance, it is time to take these models to the next level. Prospective external validation is a first necessary step before prospective interventional trials can demonstrate clinical impact. Artificial intelligence models have a huge potential in the prevention of AKI, and developing novel treatments.

### Acknowledgements

*None.*

### Financial support and sponsorship

*G.D.V. is funded by the Flemish Government [Research Foundation – Flanders (FWO grant 170719N)]. G.M. is funded by the Flemish Government [Research Foundation – Flanders (FWO)], as Senior Clinical Researcher (2012–2017: 1843113N and 2017 –2022: 1843118N) and has project funding from the KU Leuven [KU Leuven C2 project (C24/17/072): A Neuromonitor for the 21st century], as well as from the Belgian Healthcare Knowledge Centre (KCE) [COV201003: Donated antibodies working against nCoV (DAWN-Plasma)].*

### Conflicts of interest

*K.K. collaborates with Philips Research North America without a financial relationship. The authors declare no other conflicts of interest.*

## REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Hoste EAJ, Bagshaw SM, Bellomo R, *et al.* Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. Intensive Care Med 2015; 41:1411–1423.
2. Coca SG, Singanamala S, Parikh CR. Chronic kidney disease after acute kidney injury: a systematic review and meta-analysis. Kidney Int 2012; 81:442–448.

3. Silver S, Chertow G. Economic consequences of acute kidney injury. Nephron 2017; 137:297–301.
AQ8 4. Bellomo R, Ronco C, Kellum JA, *et al.* Acute renal failure – definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. Crit Care 2004; 8:R204–R212.
5. Mehta RL, Kellum JA, Shah SV, *et al.* Acute kidney injury network: report of an initiative to improve outcomes in acute kidney injury. Crit Care 2007; 11:1–8.
6. Kellum JA, Lameire N, Aspelin P, *et al.* Kidney disease: improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. Kidney Int Suppl 2012; 2:1–138.
7. Ostermann M, Pani A. Patient selection and timing of continuous renal replacement therapy. Blood Purif 2016; 42:224–237.
8. Kashani K, Rosner MH, Ostermann M. Creatinine: from physiology to clinical application. Eur J Intern Med 2020; 72:9–14.
9. Kashani KB. Automated acute kidney injury alerts. Kidney Int 2018; 94:484 –490.
10. Morris AH. Human cognitive limitations: broad, consistent, clinical application of physiological principles will require decision support. Ann Am Thorac Soc 2018; 15:S53–S56.
11. Gutierrez G. Artificial intelligence in the intensive care unit. Crit Care 2020; 24:1–9.
12. Moons KGM, Altman DG, Reitsma JB, *et al.* Transparent Reporting of a multivariable prediction model for individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015; 162:W1–W73.
13. Ahmed A, Vairavan S, Akhoundi A, *et al.* Development and validation of electronic surveillance tool for acute kidney injury: a retrospective analysis. J Crit Care 2015; 30:988–993.
14. Malhotra R, Kashani KB, Macedo E, *et al.* A risk prediction score for acute kidney injury in the intensive care unit. Nephrol Dial Transplant 2017; 32:814 –822.
15. Flechet M, Güiza F, Schetz M, *et al.* AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated lipocalin. Intensive Care Med 2017; 43:764–773.
16. Mohamadlou H, Lynn-Palevsky A, Barton C, *et al.* Prediction of acute kidney injury with a machine learning algorithm using electronic health record data. Can J Kidney Health Dis 2018; 5:2054358118776326.
17. Koyner JL, Carey KA, Edelson DP, Churpek MM. The development of a
♦ machine learning inpatient acute kidney injury prediction model. Crit Care Med 2018; 46:1070–1077.
The model makes a daily new assessment of the risk to develop acute kidney injury (AKI) in all hospitalized patients and reports the model accuracy in ICU patients separately.
18. Zimmerman LP, Reyfman PA, Smith ADR, *et al.* Early prediction of acute kidney injury following ICU admission using a multivariate panel of physiological measurements. BMC Med Inform Decis Mak 2019; 19(Suppl 1):16.
19. Chiofolo C, Chbat N, Ghosh E, *et al.* Automated continuous acute kidney
♦ injury prediction and surveillance: a random forest model. Mayo Clin Proc 2019; 94:783–792.
The model gives a new prediction for AKI in critically ill patients every 15 min.
20. Parreco J, Soe-Lin H, Parks JJ, *et al.* Comparing machine learning algorithms for predicting acute kidney injury. Am Surg 2019; 85:725–729.
21. Tomašev N, Glorot X, Rae JW, *et al.* A clinically applicable approach to continuous
&& prediction of future acute kidney injury. Nature 2019; 572:116–119.
The model is built on a very large database and continuously predicts AKI in hospitalized patients.
22. Tran NK, Sen S, Palmieri TL, *et al.* Artificial intelligence and machine learning for predicting acute kidney injury in severely burned patients: a proof of concept. Burns 2019; 45:1350–1358.
23. Rashidi HH, Sen S, Palmieri TL, *et al.* Early recognition of burn- and trauma-related acute kidney injury: a pilot comparison of machine learning techniques. Sci Rep 2020; 10:1–9.
24. Flechet M, Falini S, Bonetti C, *et al.* Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. Crit Care 2019; 23:1–10.
25. Ostermann M, Cennamo A, Meersch M, Kunst G. A narrative review of the impact of surgery and anaesthesia on acute kidney injury. Anaesthesia 2020; 75(S1):e121–e133.
26. Hobson C, Lysak N, Humber M, *et al.* Epidemiology, outcomes and management of acute kidney injury in the vascular surgery patient. J Vasc Surg 2019; 68:916 –928.
AQ9 27. Nadim MK, Forni LG, Bihorac A, *et al.* Cardiac and vascular surgery-associated acute kidney injury: the 20th International Consensus Conference of 'the ADQI (Acute Disease Quality Initiative) Group. J Am Heart Assoc 2018; 7:.
28. Thottakkara P, Ozrazgat-Baslanti T, Hupf BB, *et al.* Application of machine learning techniques to high-dimensional clinical data to forecast postoperative complications. PLoS One 2016; 11:1–19.
29. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, *et al.* MySurgeryRisk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg 2019; 269:652–662.
30. Brennan M, Puri S, Ozrazgat-Baslanti T, *et al.* Comparing clinical judgment with the MySurgeryRisk algorithm for preoperative risk assessment: a pilot usability study. Surgery 2019; 165:1035–1045.
31. Adhikari L, Ozrazgat-Baslanti T, Ruppert M, *et al.* Improved predictive models for acute kidney injury with IDEA: intraoperative data embedded analytics. PLoS One 2019; 14:1–26.
32. Lei VJ, Luong TB, Shan E, *et al.* Risk stratification for postoperative acute kidney injury in major noncardiac surgery using preoperative and intraoperative data. JAMA Netw Open 2019; 2:e1916921.
33. Lee HC, Yoon S, Yang SM, *et al.* Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. J Clin Med 2018; 7:428.
34. Cronin RM, VanHouten JP, Siew ED, *et al.* National Veterans Health Administration inpatient risk stratification models for hospital-acquired acute kidney injury. J Am Med Inform Assoc 2015; 22:1054–1071.
35. Saly D, Yang A, Triebwasser C, *et al.* Approaches to predicting outcomes in patients with acute kidney injury. PLoS One 2017; 12:1–12.
36. Garzotto F, Ostermann M, Martín-Langerwerf D, *et al.* The Dose Response Multicentre Investigation on Fluid Assessment (DoReMIFA) in critically ill patients. Crit Care 2016; 20:1–14.
37. Zhang J, Crichton S, Dixon A, *et al.* Cumulative fluid accumulation is associated with the development of acute kidney injury and nonrecovery of renal function: a retrospective analysis. Crit Care 2019; 23:1–10.
38. Hjortrup PB, Haase N, Bundgaard H, *et al.* Restricting volumes of resuscitation fluid in adults with septic shock after initial management: the CLASSIC randomised, parallel-group, multicentre feasibility trial. Intensive Care Med 2016; 42:1695–1705.
39. Myles PS, Bellomo R, Corcoran T, *et al.* Restrictive versus liberal fluid therapy for major abdominal surgery. N Engl J Med 2018; 378:2263–2274.
40. Zhang Z, Ho KM, Hong Y. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care. Crit Care 2019; 23:1–10.
41. Lin K, Hu Y, Kong G. Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model. Int J Med Inform 2019; 125:55 –61.
42. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019; 380:1347–1358.
43. Hunter JS. Enhancing Friedman's 'fundamental theorem of biomedical informatics'. J Am Med Inform Assoc 2010; 17:112–113.
44. Sutherland SM, Chawla LS, Kane-Gill SL, *et al.* Utilizing electronic health records to predict acute kidney injury risk and outcomes: workgroup statements from the 15th ADQI Consensus Conference. Can J Kidney Health Dis 2016; 3:1–14.
45. Kappen TH, van Klei WA, van Wolfswinkel L, *et al.* Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res 2018; 2:1–11.
46. Kashani K, Dalili N, Carter R, *et al.* Using clinical decision support systems for acute kidney injury pragmatic trials. J Transl Crit Care Med 2019; 1:28.
47. Davis SE, Lasko TA, Chen G, *et al.* Calibration drift in regression and machine learning models for acute kidney injury. J Am Med Inform Assoc 2017; 24:1052– 1061.
48. Beam AL, Kohane IS. Big data and machine learning in healthcare. JAMA 2018; 319:1317–1318.
49. Kappen TH, Vergouwe Y, Van Klei WA, *et al.* Adaptation of clinical prediction models for application in local settings. Med Decis Making 2012; 32:1 – 10.
50. Haase M, Kellum JA, Ronco C. Subclinical AKI – an emerging syndrome with important consequences. Nat Rev Nephrol 2012; 8:735–739.
51. Carsetti A, Cecconi M, Rhodes A. Fluid bolus therapy: monitoring and predicting fluid responsiveness. Curr Opin Crit Care 2015; 21:388–394.
52. Kappen TH, Van Loon K, Kappen MAM, *et al.* Barriers and facilitators perceived by physicians when using prediction models in practice. J Clin Epidemiol 2016; 70:136–145.
53. Winters BD, Cvach MM, Bonafide CP, *et al.* Technological distractions (part 2): a summary of approaches to manage clinical alarms with intent to reduce alarm fatigue. Crit Care Med 2018; 46:130–137.
54. Hoste EAJ, Kashani K, Gibney N, *et al.* Impact of electronic-alerting of acute kidney injury: workgroup statements from the 15th ADQI Consensus Conference. Can J Kidney Health Dis 2016; 3:1–9.
55. Kashani K, Al-Khafaji A, Ardiles T, *et al.* Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. Crit Care 2013; 17:1–12.
56. Hoste EA, Vaara ST, De Loor J, *et al.* Urinary cell cycle arrest biomarkers and chitinase 3-like protein 1 (CHI3L1) to detect acute kidney injury in the critically ill: a post hoc laboratory analysis on the FINNAKI cohort. Crit Care 2020; 24:1–10.

MCC

Manuscript No. MCC260604

Current Opinion in Critical Care
Typeset by Thomson Digital
for Wolters Kluwer

Dear Author,

During the preparation of your manuscript for typesetting, some queries have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin as neatly as possible or compile them as a separate list. This form should then be returned with your marked proof/list of corrections to the Production Editor.

## QUERIES: to be answered by AUTHOR/EDITOR?

| QUERY NO. | QUERY DETAILS | RESPONSE |
|---|---|---|
| <AQ1> | As per style, the short title/running head can have a maximum of 65 characters including spaces and author names, and abbreviations/acronyms only as exceptions. Please check the suggested short title, 'Acute kidney injury in the ICU' for appropriateness. | |
| <AQ2> | Please confirm whether surnames/family names (red) have been identified correctly in the author byline. | |
| <AQ3> | Please check the affiliations for correctness. | |
| <AQ4> | Affiliation 'b' has been split into two different affiliations. Please check, and correct if necessary. | |
| <AQ5> | Please check the current corresponding author information for correctness. | |
| <AQ6> | Please check the keywords for correctness. | |
| <AQ7> | Please provide full form for 'ARDS' as per style. | |
| <AQ8> | Refs. [4,7,16,18] have been updated using PubMed. Please check for correctness of information. | |
| <AQ9> | In Ref. [27], page range is not available in PubMed. Please update as per style. | |
| <AQ10> | Please check the layout of Tables 1–3 and correct if necessary. | |