

Internal and External Validation of a Machine Learning-assisted Predicting Models for Mechanical Ventilation-associated Severe Acute Kidney Injury

Cong Feng (✉ fc194@163.com)

Research

Keywords: Severe acute kidney injury, Mechanical ventilation, Machine learning, Prediction model

Posted Date: January 28th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1305991/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

The preventative and therapeutic strategies of mechanical ventilation (MV)-associated severe AKI are still limited. We developed clinical prediction models for early warning the occurrence of severe AKI in the first week of ICU during the initiation of MV.

Methods

Retrospective analysis of a large ICU database (MIMIC-IV). Data collection were based on the clinical information available on ICU admission and during the first 12 hours of MV. Predictors were selected successively using univariable and multivariable analysis. Two machine learning algorithms were compared for model development. The primary outcome was predicting the development of AKI stage 2 or 3 (AKI-23) and AKI stage 3 (AKI-3) in the first week of ICU stay after the initiation 12 hours of MV. The developed models were further external validated in another multi-center ICU database (eICU) and evaluated in different subpopulations of patients.

Results

Models were developed in the development cohort (MIMIC-IV: 2008-2016 n=3,986) and random forest performed better than logistic regression. In the internal (MIMIC-IV: 2017-2019, n=1,210) and external (eICU, n=1,494) validation cohort, the incidences of AKI-23 were 154 (12.7%) and 119 (8.0%), the AUCs were 0.78 (95% CI 0.74-0.82) and 0.80 (95% CI 0.76-0.84), respectively, while AKI-3 had the incidences were 81 (6.7%) and 67 (4.5%), the AUCs were 0.81 (95% CI 0.76-0.87) and 0.80 (95% CI 0.73-0.86), respectively.

Conclusions

MV-associated severe AKI can be predicted early with models driven by machine learning techniques based on routinely clinical information. The validated models are available at:
https://apoet.shinyapps.io/mv_aki_2021_V2.0/

Background

Mechanical ventilation (MV) is a critically life-saving intervention which is widely required in intensive care unit (ICU). Moreover, it is also regarded as one of the most important risk factors for acute kidney injury (AKI) in critically ill patients, which could increase the odds of AKI nearly threefold(1). Severe AKI, double serum creatinine or the need of acute dialysis, is independently associated with an up to five-fold higher risk of death (2, 3). The patients with combined MV and severe AKI have a higher mortality rate of

approximately 60-80%(3–7). Up to now, the therapeutic strategies for MV-induced severe AKI are still limited (6), which partly due to the delays in the identification of severe AKI on the basis of elevation in serum creatinine (SCr) 2-3 days after the initial injury. With an effective treatment of AKI lacking, prevention becomes greatly critical since it would greatly improves the poor outcomes of AKI in the ICU (8). Before a detectable rise in SCr, early detection of patients at increased risk of MV-induced severe AKI, would provides intensivists a critical time window in early evaluation and better informed decision making to change the current intervention strategies utilized in a MV-receiving patient (9, 10).

Currently, different pathophysiological and underlying processes were explored in experimental studies to explain the interactions between MV and AKI(11, 12). But there is no an ideal method that could provide the early information of the odds of severe AKI to clinicians after the initiation of MV(13, 14). The increasing secondary analysis of large electronic data of severely ill patients provides a new opportunity and a major step forward to predict the occurrence of severe AKI in ICU (15–20). It has been extensively suggested that such prediction model based on clinical data and machine learning would predict much earlier, be more applicable and cost-effective than currently delayed diagnostic criteria and limited biomarkers. We therefore hypothesized that machine learning might be a new method for early prediction of severe AKI specifically in mechanically ventilated patients who should be paid more attention and cost too much medical resources differently from non-mechanically ventilated patients.

The aim of this study was the development of the early clinical prediction models for the occurrence of severe AKI in the first week of ICU stay after the initiation of MV, which were based on the large ICU database with general population background and machine learning algorithms. The developed models would be further external validated in another large ICU database and different subpopulations.

Methods

Study design and cohorts

The prediction models were developed and internal validated by the retrospective analysis of a single-center publicly available ICU database (MIMIC-IV,v0.4) which contained 76540 ICU admissions during 2008–2019 in 6 different ICUs from a teaching hospital (Beth Israel Deaconess Medical Center, BIDMC; Boston, Massachusetts)(21-23). We divided the MIMIC-IV into development (2008–2016) and validation (2017–2019) cohorts for model development and internal validation. External validation was performed in a multi-center ICU database (eICU, v1.2), which included 459 ICUs from 58 hospitals during 2003–2016 in the United States.

Both databases contained most of the routinely clinical patients' data including demographics, vital signs, laboratory tests, and other information of diagnosis, medications, and procedures. Data collection for the database was passive and had no impact on patients' safety. The databases were de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. The secondary analysis of both databases was approved by the team of the Laboratory for Computational

Physiology from the Massachusetts Institute of Technology (MIT-LCP) who contributed and maintained the MIMIC-IV and eICU databases.

All patients in the MIMIC-IV and eICU were included in this study and only data of the first ICU admission of the first hospitalization were used (24). The inclusion criteria were as follows: 1) Age \geq 18 years old; 2) ICU-stay for at least 24 consecutive hours; 3) Invasive MV, including ventilation through a endotracheal intubation, started on or after ICU admission and for at least 12 consecutive hours. The exclusion criteria were as follows: 1) The clinical history of the end-stage renal disease (ESRD); 2) The baseline SCr \geq 4 mg/dL; 3) Need for dialysis or renal replacement therapy (RRT) within 24 hour after ICU admission; 4) The measurement of SCr was inadequate to stage AKI (SCr missing on day 1 or missing for more than 48 hours); 5) Patients with insufficient data collection.

Severe AKI definitions

The AKI was staged daily during ICU stay based on the SCr criteria of "Kidney disease: Improved Global Outcomes" (KDIGO) (25). Severe AKI was defined as AKI stage 2 or 3 according to 48-hour and 7-day SCr criteria of KDIGO (25). The mean SCr measurements prior to ICU admission in databases were considered as the baseline SCr. The SCr measured on ICU admission was used as an imputation value in patients with missing baseline SCr (26). The criteria of acute dialysis or RRT were excluded in our study since our models would be further evaluated in predicting the requirement for dialysis or RRT during ICU stay.

Prediction tasks performed

The primary prediction task of this study was predicting the development of severe AKI in the first week of ICU stay after the initiation 12 hours of MV. The details of two primary outcomes were elaborated as AKI-23 and AKI-3:

- 1) AKI-23: Prediction of the occurrence of AKI stage 2 or 3 during the first week of ICU stay after the initiation 12 hours of MV.
- 2) AKI-3: Prediction of the first occurrence of AKI stage 3 during the first week of ICU stay after the initiation 12 hours of MV.

Prediction models

For each prediction task, one model was developed based on the clinical information, available on ICU admission and the initiation 12 hours after MV. Models were developed and internal validated using the logistic regression and random forest machine-learning algorithms, in order to choose the better models

for further investigation. The logistic regression and random forest machine-learning algorithms were performed by R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) and the R package “randomForest” (20).

Predictor selection

For each patient, most of the routinely available ICU data known to be associated with MV and AKI was used candidate predictors according to literature review and availability in database (17, 26). In the development cohort, the potential predictors entered into the selection process. The final set of predictor variables was successively determined by univariable and multivariable analyses for each predicting model. The variables were selected as the final predictors when they showed a strong association with severe AKI when the statistical difference at the 0.10 level was observed in both univariable and multivariable analyses. The univariable and multivariable analyses were performed by R version 3.5.1.

Model development and validation

The selected variables to be used as predictors for each model, as well as model development was performed in the development cohort only. Model performance and stability were validated via 10-fold cross-validation in the development cohort. The performance of the different models was evaluated in the internal validation cohort to choose the better machine-learning algorithm. Then, models were further validated in the external cohort. Models were subsequently evaluated separately in cardiac, respiratory, nervous and septic patients only, and for prediction of the requirement of dialysis or renal replacement therapy (RRT) during ICU stay. Model development, internal and external validation were performed by R version 3.5.1.

Performance evaluation and statistical analysis

The continuous variables were expressed as median and interquartile range (IQR). The significance of the differences between two groups was determined using the Mann-Whitney U test or Student's t-test, as appropriate. Categorical variables were expressed as absolute (n) and relative (%) frequency and analyzed by chi-square test or Fisher's exact test, as appropriate. The multivariable analysis was assessed using the logistic regression. A two-sided α of less than 0.05 was considered statistically significant. To compare performance criteria of models in the same patient subset, the difference and corresponding bootstrap confidence intervals (CIs) were computed. If the 95% CI excluded 0, the difference was declared statistically significant at the 0.05 level.

To assess the performance of the developed models, the evaluations for discrimination and calibration were reported (20, 27). Discrimination referred to how well the predictions allow discriminating positive or negative outcomes, which was commonly evaluated with the receiver operator characteristic (ROC) curve, and quantified with the area under the ROC curve (AUC).

The additional common measures of discriminability were also reported including sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) at the classification thresholds that maximized sensitivity and specificity in the development cohort. The classification threshold referred to the cutoff for the predicted probabilities above which the patient would be considered 'positive' for the outcome. We have chosen the cutoff that result in the highest sensitivity and specificity in a population with similar characteristics (as determined for all patients in the development cohort) in the reported tables and as the default in the online calculator. This cutoff was selected equivalent to the optimal point on the ROC curve based on the maximum value of the Youden-index. All statistical analysis was performed by R version 3.5.1.

Results

Study population: development, internal and external validation cohorts

A total of 5,196 patients of MIMIC-IV (2008-2016: 3,986; 2017-2019: 1,210) and 1,494 patients of eICU were included according to the inclusion and exclusion criteria presented in Figure 1 and Supplemental Figure 1. Within the first week of ICU stay, 421 (10.6%), 154 (12.7%) and 119 (8.0%) MV patients developed AKI-23 in the development, internal and external validation cohorts, respectively. There were 2,069 (51.9%), 253 (20.9%) and 411 (27.5%) septic patients, and 785 (19.7%), 306 (25.3%) and 463 (31.0%) cardiac patients in the development, internal and external validation cohorts, respectively. Meanwhile, the median hour from MV to AKI-23 was 52.6 (37.3-122.8), 46.5 (31.5-107.0) and 44.0 (22.0-85.5) in the development, internal and external validation cohorts, respectively. The SCr baseline was 1.0 (0.8-1.4) mg/dL in all cohorts. The need for dialysis or RRT during ICU stay was 203 (5.1%) in the development cohort, 104 (8.6%) in the internal validation cohort, and 37 (2.5%) in the external validation cohort. Other detail information of each cohort was shown in Table 1 and Supplemental Table 1. The information about 7,458 patients excluded for insufficient data in the MIMIC-IV was provided in the Supplemental Table 2.

Variable selection

We ended up extracting a set of 57 candidate variables came from the time before, upon ICU admission, and after MV, including demographic information, primary diagnosis, comorbidities, vital signs, ventilator parameters, additional hemodynamic support, and laboratory values from the databases. Details of the candidate predictors were reported in Supplemental Table 3. The univariable association with AKI-23 and AKI-3 in the development cohort was listed in the Supplemental Table 4-5. The selected predictors based

on further multivariable analysis for AKI-23 and AKI-3 model development were shown in Supplemental Table 6-7. In the final, there are 12 and 8 variables were selected as the predictors for AKI-23 and AKI-3 model development, respectively.

Development cohort model performance

In the development cohort, the performances of AKI-23 and AKI-3 models based on random forest were better than those based on logistic regression. The performance of AKI-23 model was with the AUCs of 0.77 (95% CI 0.69-0.84) for logistic regression and 0.82 (95% CI 0.76-0.88) for random forest. Meanwhile, the AUCs of AKI-3 model were 0.79 (95% CI 0.70-0.88) and 0.89 (95% CI 0.84-0.95) for logistic regression and random forest, respectively. Based on the maximum Youden index, the PPVs of AKI-23 model were 0.25 (95% CI 0.20-0.36) for logistic regression and 0.39 (95% CI 0.17-0.53) for random forest. The PPVs of AKI-3 model were 0.21 (95% CI 0.11-0.32) and 0.27 (95% CI 0.14-0.40) for logistic regression and random forest, respectively. The other discrimination results of the clinical prediction models for AKI-23 and AKI-3 in the development cohort based on two machine-learning algorithms were shown in Table 2. The random forest was selected as the final machine-learning algorithm for model development. To improve specificity or PPV, other classification thresholds within the range of clinical usefulness could be chosen (Supplemental Table 8).

Internal validation cohort model performance

The performances for AKI-23 and AKI-3 models developed by random forest in the internal validation cohort were reported in Table 3. The discrimination was shown with AUCs of 0.78 (95% CI 0.74-0.82) and 0.81 (95% CI 0.76-0.87) for AKI-23 and AKI-3, respectively. The classification thresholds identified in the development cohort were 23.5% and 11.9% for AKI-23 and AKI-3, respectively, remained robust and resulted in similar sensitivities, specificities, PPV, and NPV in the internal validation cohort.

The calibration of the models

AKI-23 and AKI-3 models were well calibrated with respective calibration slopes of 1.13 and 1.08 for development cohort, 1.04 and 0.90 for internal validation cohort. The calibration-in-the-large was close to 0 and calibration curve was close to the diagonal (Supplemental Figure 2).

External validation cohort model performance

Table 4 reported the performances of the clinical prediction models for AKI-23 and AKI-3 in the external validation cohort. The discrimination results, based on the classification thresholds identified in the development cohort, had the AUC of 0.80 (95% CI 0.76-0.84), sensitivity of 0.57 (95% CI 0.48-0.66), specificity of 0.83 (95% CI 0.81-0.85), PPV of 0.23 (95% CI 0.20-0.27) and NPV of 0.96 (95% CI 0.95-0.97) for AKI-23, and AUC of 0.80 (95% CI 0.73-0.86), sensitivity of 0.67 (95% CI 0.55-0.78), specificity of 0.85 (95% CI 0.83-0.87), PPV of 0.18 (95% CI 0.15-0.21), and NPV of 0.98 (95% CI 0.98-0.99) for AKI-3.

Model performance in the subpopulations

The AKI-23 and AKI-3 models performed well in all the subpopulations of cardiac patients, respiratory patients, septic patients and nervous patients from the internal and external validation cohorts. In the internal validation cohort, the AUCs were between 0.68 and 0.85 for AKI-23, and between 0.70 and 0.77 for AKI-3. Meanwhile, in the external validation cohort, the AUCs were between 0.73 and 0.85, and between 0.74 and 0.77 for AKI-23 and AKI-3, respectively. The other detailed results were showed in Supplemental Table 9-10.

Model performance of major clinical outcome

All AKI-23 and AKI-3 models performed well at predicting the requirement of dialysis or RRT with AUCs between 0.76 and 0.77 in the internal validation cohort, and 0.76 and 0.78 in the external validation cohort. The other detailed results were illustrated in Supplemental Table 11.

Online MV-associated severe AKI predicting models

We made the fine-tuned models MV-associated AKI-23 and AKI-3 publicly available through our online portal at https://apoet.shinyapps.io/mv_aki_2021_V2

Discussion

The inverse relationship between MV and kidney function was first observed several decades ago and becoming a critically global challenge because of the greatly limited management (1, 28, 29). One of the reasons of lacking preventative and therapeutic strategies is the short of a desirable method offering the early-warning valuable information in evaluating the risk of AKI after the initiation of MV (17, 30). Since it was highlighted that severe AKI had strong association with poor clinical outcomes(2, 3, 31), we developed and validated the clinical prediction models for early MV-associated AKI-23 and AKI-3 prognostication based on large databases in this study. Prediction of severe AKI stage is relevant to clinical practice and research, as it allows for risk stratification. Additionally, in clinical trials, it might aid

in stratification or heterogeneity reduction. Higher-risk patients require special attention and might benefit from an earlier preventive management.

Clinical prediction for prognostication or decision-support in critical care about MV or AKI based on machine learning has been extensively studied and formed many cost-effective models (20, 24, 26, 30–43). However, the clinical prediction models for the development of severe AKI after the initiation of MV are scarce. Our results clearly demonstrated that the development of AKI-23 and AKI-3 after the initiation 12 hours of MV during the first week of ICU stay, as defined by the KDIGO SCr criteria, could be accurately predicted based only on clinical information. Although fewer predictor variables were selected, the developed models performed well according to evaluation criteria in internal and external validation(20), and even in different subpopulations of patients. The models also significantly outperformed in predicting the associated requirement of dialysis or RRT during ICU stay.

It is well known that the invasive MV could be as a risk factor for AKI in the critically ill patients based on a systematic review and meta-analysis (1), yet the relationship between MV-related parameters and AKI is not clear. The meta-analysis, including many observations with different levels of tidal volume and PEEP, found no association between MV-related parameters and AKI(1). Similarly, another retrospective study found no significant association with any early or late AKI(12). In this study, based on multivariable analysis, we could not found the strong association with AKI-23 or AKI-3 either. However another study, investigating the patients with acute respiratory distress syndrome in MIMIC-III database, has found that PEEP is the only respiratory-related variable with a direct causal association in severe AKI(44). Although PEEP may reduce cardiac output and increase central venous pressure, which could diminish renal blood flow, free water clearance, or the glomerular filtration rate, it should be further evaluated and validated in a multivariate model including other potentially correlated variables(11). Therefore, the relationship between MV-related parameters and severe AKI still need further investigation.

For non-MV-related risk factors, in accordance with previous study(12), the platelet count in the initiation of MV showed strong association with AKI-23 and AKI-3. Moreover, we found that the lowest /diastolic blood pressure, serum phosphorus and magnesium in the first 12 hours after MV also had multivariable association with AKI-23 and AKI-3. It was suggested that these risk factors might related to the underlying patho-physiological mechanisms of MV-kidney interactions, which could help to conduct further investigations(45). Meanwhile, they also could be used as the potential clinical indicators for early risk assessment of severe AKI in MV-receiving patients.

We made the developed and validated AKI-23 and AKI-3 models publicly available as an online prognostic calculator. Since early detection of MV-receiving patients at high risk of severe AKI, before a detectable rise in SCr, would bring a critical time window, clinicians could halt or reverse the ongoing renal injury in time(46). Furthermore, different classification thresholds are likely required to study populations at different risks for MV-associated severe AKI. For example, a higher threshold would be considered when studying septic patients than for nervous patients, which may have a better specificity and positive

predictive value. In this study, we also provided multiple classification thresholds in the reported tables and in the online calculator for different clinical and research implications.

This study has several strengths. The first is its use of a large dataset of diverse patient populations who received MV from six ICUs (21). Second, patients who received MV for a minimum of 12h were included, which would constitute much earlier warning window than other studies to identify the high risk of developing the severe AKI (44, 47). Third, as the included patients under mix modes of MV came from multiple ICUs where ventilation practices vary, our models might be more applicable (44). Fourth, a development, internal validation and external validation approach was followed, with robust machine learning algorithm compared and selected for model development and further validation in different subpopulations. Fifth, we integrated most of the routinely available ICU data as the candidate predictors for selecting, which successively illustrated by univariable and multivariable association analysis (20). Sixth, the models used the least number of predictors that are likely available in most ICUs, which improve their likelihood of worldwide use and generalization. Seventh, the need for dialysis or RRT during ICU stay could also be predicted well base on the developed AKI-23 and AKI-3 models. Last, the developed and validated AKI-23 and AKI-3 models are publicly available as an online early warning tool. It provides an useful platform for further model validation, future biomarker, MV parameters, MV-associated severe AKI prediction studies, and complements existing risk evaluation scores in use at the bedside and/or in limited resource settings.

This study still has some limitations. First, the development cohort came from single clinical center while multi-center studies might be more convincing. However, the developed models were validated well in an independent multi-center dataset. Second, we excluded patients whose measurement of SCr was inadequate to stage AKI and with insufficient data collection, so our models do not apply to these patients and limited their generalization. Third, some routinely clinical variables, such as height, fluid intake and other lab values, were not included as the candidate variables for predictor selection and model development due to the high missing ratio in the databases, while these features might improve the performance for early warning. Fourth, the causal association between MV parameters and severe AKI could not be demonstrated solely based on the machine learning algorithms. Fifth, as only the KDIGO SCr criteria were used, patients developing severe AKI according to KDIGO urine output criteria could be missed. Nevertheless, this is in line with several other previously published AKI risk scores and the models predicted well the related and clinically relevant outcomes of dialysis or RRT requirement at ICU discharge. Sixth, study design in this work relied on retrospective data investigation which may cause missing some important information compared to a prospective study. Besides, results about the impact analysis between model prediction and patients' outcomes could not be obtained. Seventh, the modest model performance might be contributed by the imperfect classification, since the AKI phenotype has not truly been established when MV associated with AKI. Last, several steps still need to be taken before such models could be translated into clinical practice. Therefore, these models also need further external validation in other independent datasets and prospective validation in a real-world clinical setting to demonstrate that preventive actions triggered by model outputs meaningfully alter patient outcomes, patient experiences, or care processes.

Conclusions

We have shown that the development of AKI-23 and AKI-3 after the initiation 12 hours of MV during the first week of ICU stay, as defined by the KDIGO SCr criteria, can be accurately predicted by machine learning techniques based on clinical information. These developed clinical prediction models are available online to facilitate further researches, which is freely accessible through https://apoet.shinyapps.io/mv_aki_2021_V2.

Abbreviations

MV
mechanical ventilation

AKI
acute kidney injury

AKI-23
AKI stage 2 or 3

AKI-3
AKI stage 3

ICU
intensive care unit

SCr
serum creatinine

HIPAA
Health Insurance Portability and Accountability Act

MIT-LCP
Laboratory for Computational Physiology from the Massachusetts Institute of Technology

ESRD
end-stage renal disease

RRT
renal replacement therapy

KDIGO
Kidney disease:Improved Global Outcomes

IQR
interquartile range

CIs
confidence intervals

ROC
receiver operator characteristic

AUC
area under the ROC curve

PPV

positive predictive value

NPV

negative predictive value

Declarations

Ethics approval and consent to participate

Data collection was passive and had no impact on patient safety. The data set was de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Studies performed on de-identified data constitute non-human subject research, thus no institutional or ethical approvals were required for this study.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and during the current study are available in the MIMIC and eICU repository, <https://physionet.org/>.

Competing interests

The authors declare that they have no competing interests.

Funding

There is no financial funding or interest to report.

Author's Contributions

CF designed the study, conducted the data collection, data analysis, data interpretation, and wrote the manuscript. SH conducted the data analysis, data interpretation, conducted the online early warning calculator, developed the website, and wrote the manuscript. LC and YB conducted the data interpretation and wrote the manuscript. LW and XZ conducted the data interpretation and reviewed the manuscript.

Acknowledgements:

We thank the team of the Laboratory for Computational Physiology from the Massachusetts Institute of Technology (MIT-LCP) who keep the MIMIC databases. We also thank the support from the National Natural Science Fund (no. 82072200) and from the Chinese PLA General Hospital (2020-YQPY-002, 2019XXJSYX20, 2019XXMBD-016).

References

1. van den Akker JP, Egal M, Groeneveld AB. Invasive mechanical ventilation as a risk factor for acute kidney injury in the critically ill: a systematic review and meta-analysis. *Crit Care (London England)*. 2013;17(3):R98.
2. Hoste EA, Bagshaw SM, Bellomo R, Cely CM, Colman R, Cruz DN, et al. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. *Intensive Care Med*. 2015;41(8):1411–23.
3. Nisula S, Kaukonen KM, Vaara ST, Korhonen AM, Poukkanen M, Karlsson S, et al. Incidence, risk factors and 90-day mortality of patients with acute kidney injury in Finnish intensive care units: the FINNAKI study. *Intensive Care Med*. 2013;39(3):420–8.
4. Bellomo R, Kellum JA, Ronco C. Acute kidney injury. *Lancet*. 2012;380(9843):756–66.
5. Druml W, Lenz K, Laggner AN. Our paper 20 years later: from acute renal failure to acute kidney injury—the metamorphosis of a syndrome. *Intensive care medicine*. 2015;41(11):1941–9.
6. Hepokoski ML, Malhotra A, Singh P, Crotty Alexander LE. Ventilator-Induced Kidney Injury: Are Novel Biomarkers the Key to Prevention? *Nephron*. 2018;140(2):90–3.
7. Lombardi R, Nin N, Lorente JA, Frutos-Vivar F, Ferguson ND, Hurtado J, et al. An assessment of the Acute Kidney Injury Network creatinine-based criteria in patients submitted to mechanical ventilation. *Clin J Am Soc Nephrol*. 2011;6(7):1547–55.
8. Schetz M, Schneider A. Focus on acute kidney injury. *Intensive Care Med*. 2017;43(9):1421–3.
9. Pickkers P, Ostermann M, Joannidis M, Zarbock A, Hoste E, Bellomo R, et al. The intensive care medicine agenda on acute kidney injury. *Intensive Care Med*. 2017;43(9):1198–209.
10. Hepokoski M, Owens RL, Malhotra A, Beitler JR. Mechanical ventilation in acute respiratory distress syndrome at ATS 2016: the search for a patient-specific strategy. *Journal of thoracic disease*. 2016;8(Suppl 7):550-2.
11. Darmon M, Legrand M, Terzi N. Understanding the kidney during acute respiratory failure. *Intensive care medicine*. 2017;43(8):1144–7.
12. Lombardi R, Nin N, Peñuelas O, Ferreiro A, Rios F, Marin MC, et al. Acute Kidney Injury in Mechanically Ventilated Patients: The Risk Factor Profile Depends on the Timing of Aki Onset. *Shock (Augusta Ga)*. 2017;48(4):411–7.
13. Prowle JR. Measurement of AKI biomarkers in the ICU: still striving for appropriate clinical indications. *Intensive care medicine*. 2015;41(3):541–3.
14. Bell M, Larsson A, Venge P, Bellomo R, Mårtensson J. Assessment of cell-cycle arrest biomarkers to predict early and delayed acute kidney injury. *Disease markers*. 2015;2015:158658.
15. Hodgson LE, Selby N, Huang TM, Forni LG. The Role of Risk Prediction Models in Prevention and Management of AKI. *Semin Nephrol*. 2019;39(5):421–30.
16. Wilson FP. Machine Learning to Predict Acute Kidney Injury. *American journal of kidney diseases: the official journal of the National Kidney Foundation*. 2020;75(6):965–7.

17. Hodgson LE, Sarnowski A, Roderick PJ, Dimitrov BD, Venn RM, Forni LG. Systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations. *BMJ open*. 2017;7(9):e016591.
18. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. 2019;572(7767):116–9.
19. Park S, Lee H. Acute kidney injury prediction models: current concepts and future strategies. *Curr Opin Nephrol Hypertens*. 2019;28(6):552–9.
20. Flechet M, Güiza F, Schetz M, Wouters P, Vanhorebeek I, Derese I, 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 medicine*. 2017;43(6):764–73.
21. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035.
22. Johnson AE, Stone DJ, Celi LA, Pollard TJ. The MIMIC Code Repository: enabling reproducibility in critical care research. *J Am Med Inform Assoc*. 2018;25(1):32–9.
23. Johnson ABL, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV (ver-
24. sion 0.4) 2020 [Available from: <https://www.physionet.org/>].
25. Serpa Neto A, Deliberato RO, Johnson AEW, Bos LD, Amorim P, Pereira SM, et al. Mechanical power of ventilation is associated with mortality in critically ill patients: an analysis of patients in two observational cohorts. *Intensive Care Med*. 2018;44(11):1914–22.
26. Thomas ME, Blaine C, Dawnay A, Devonald MA, Ftouh S, Laing C, et al. The definition of acute kidney injury and its use in practice. *Kidney Int*. 2015;87(1):62–73.
27. Malhotra R, Kashani KB, Macedo E, Kim J, Bouchard J, Wynn S, et al. A risk prediction score for acute kidney injury in the intensive care unit. *Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association - European Renal Association*. 2017;32(5):814–22.
28. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA*. 2017;318(14):1377–84.
29. Drury DR, Henry JP, Goodman J. The Effects of Continuous Pressure Breathing on Kidney Function. *J Clin Investig*. 1947;26(5):945–51.
30. Husain-Syed F, Slutsky AS, Ronco C. Lung-Kidney Cross-Talk in the Critically Ill Patient. *Am J Respir Crit Care Med*. 2016;194(4):402–14.
31. Darmon M, Ostermann M, Joannidis M. Predictions are difficult... especially about AKI. *Intensive Care Med*. 2017;43(6):932–4.
32. Sutherland SM, Chawla LS, Kane-Gill SL, Hsu RK, Kramer AA, Goldstein SL, et al. Utilizing electronic health records to predict acute kidney injury risk and outcomes: workgroup statements from the 15(th) ADQI Consensus Conference. *Can J Kidney Health Dis*. 2016;3:11.

33. Abul-Husn NS, Kenny EE. Personalized Medicine and the Power of Electronic Health Records. *Cell*. 2019;177(1):58–69.
34. Ghassemi M, Wu M, Hughes MC, Szolovits P, Doshi-Velez F. Predicting intervention onset in the ICU with switching state space models. *AMIA Jt Summits Transl Sci Proc*. 2017;2017:82-91.
35. Hodgson LE, Dimitrov BD, Roderick PJ, Venn R, Forni LG. Predicting AKI in emergency admissions: an external validation study of the acute kidney injury prediction score (APS). *BMJ open*. 2017;7(3):e013511.
36. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. *Ann Thorac Surg*. 2012;93(1):337–47.
37. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature medicine*. 2018;24(11):1716–20.
38. Koynert JL, Adhikari R, Edelson DP, Churpek MM. Development of a Multicenter Ward-Based AKI Prediction Model. *Clinical journal of the American Society of Nephrology: CJASN*. 2016;11(11):1935–43.
39. Mohamadlou H, Lynn-Palevsky A, Barton C, Chettipally U, Shieh L, Calvert J, et al. Prediction of Acute Kidney Injury With a Machine Learning Algorithm Using Electronic Health Record Data. *Canadian journal of kidney health disease*. 2018;5:2054358118776326.
40. Pozzoli S, Simonini M, Manunta P. Predicting acute kidney injury: current status and future challenges. *J Nephrol*. 2018;31(2):209–23.
41. Silver SA, Shah PM, Chertow GM, Harel S, Wald R, Harel Z. Risk prediction models for contrast induced nephropathy: systematic review. *BMJ*. 2015;351:h4395.
42. Smith LE, Smith DK, Blume JD, Siew ED, Billings FTt. Latent variable modeling improves AKI risk factor identification and AKI prediction compared to traditional methods. *BMC Nephrol*. 2017;18(1):55.
43. Suresh H, Hunt N, Johnson A, Anthony Celi L, Szolovits P, Ghassemi M. Clinical Intervention Prediction and Understanding using Deep Networks 2017.
44. Wilson T, Quan S, Cheema K, Zarnke K, Quinn R, de Koning L, et al. Risk prediction models for acute kidney injury following major noncardiac surgery: systematic review. *Nephrol Dial Transplant*. 2016;31(2):231–40.
45. Leite TT, Gomes CAM, Valdivia JMC, Libório AB. Respiratory parameters and acute kidney injury in acute respiratory distress syndrome: a causal inference study. *Annals of translational medicine*. 2019;7(23):742.
46. Darmon M, Schortgen F, Leon R, Moutereau S, Mayaux J, Di Marco F, et al. Impact of mild hypoxemia on renal function and renal resistive index during mechanical ventilation. *Intensive care medicine*. 2009;35(6):1031–8.
47. Yang X, Chen C, Teng S, Fu X, Zha Y, Liu H, et al. Urinary Matrix Metalloproteinase-7 Predicts Severe AKI and Poor Outcomes after Cardiac Surgery. *Journal of the American Society of Nephrology: JASN*. 2017;28(11):3373–82.

48. Lombardi R, Nin N, Penuelas O, Ferreiro A, Rios F, Marin MC, et al. Acute Kidney Injury in Mechanically Ventilated Patients: The Risk Factor Profile Depends on the Timing of Aki Onset. Shock. 2017;48(4):411–7.

Tables

Table 1. Description of the development and internal validation cohorts.

Key Characteristics	Development cohort (MIMIC-IV: 2016)	Internal validation cohort (MIMIC-IV: 2017-2019)	P-value
Demographic information:			
Final cohort (n)	3,986	1,210	-
Age, years (median, (IQR))	65.0 (52.0-77.0)	65.0 (53.0-75.0)	0.86
Weight, kg (median, (IQR))	79.8 (67.0-95.6)	80.2 (68.0-96.5)	0.31
Male gender (n(%))	2,283 (57.3%)	734 (60.7%)	<0.05
Serum creatinine baseline, mg/dL (median, (IQR))	1.0 (0.8-1.4)	1.0 (0.8-1.4)	0.40
The time from mechanical ventilation to AKI-23, hour (median, (IQR))	52.6 (37.3-122.8)	46.5 (31.5-107.0)	0.17
The time from mechanical ventilation to AKI-3, hour (median, (IQR))	58.7 (38.4-110.4)	46.6 (35.0-107.0)	0.24
The incidence within the first week (n(%)):			
AKI-123	978 (24.5%)	326 (26.9%)	0.10
AKI-23	421 (10.6%)	154 (12.7%)	<0.05
AKI-1	557 (14.0%)	172 (14.2%)	0.47
AKI-2	179 (4.5%)	73 (6.0%)	<0.05
AKI-3	242 (6.1%)	81 (6.7%)	0.87
Comorbidities (n(%)):			
Hypertension	1,814 (45.5%)	564 (46.6%)	0.52
Diabetes	1,109 (27.8%)	294 (24.3%)	<0.05
Congestive heart failure	981 (24.6%)	303 (25.0%)	0.79
Chronic pulmonary disease	588 (14.8%)	338 (27.9%)	<0.01
Chronic kidney disease	619 (15.5%)	200 (16.5%)	0.43

Chronic liver disease	165 (4.1%)	116 (9.6%)	<0.01
Primary diagnosis (n(%)):			
Sepsis, including pneumonia	2,069 (51.9%)	253 (20.9%)	<0.01
Cardiovascular	785 (19.7%)	306 (25.3%)	<0.01
Other Respiratory	686 (17.2%)	382 (31.6%)	<0.01
Neurological	250 (6.3%)	144 (11.9%)	<0.01
Others	196 (4.9%)	125 (10.3%)	<0.01
Ventilator parameters:			
The highest FIO ₂ of ventilator, % (median, (IQR))	55.0 (45.0-66.7)	50.0 (40.0-62.5)	<0.01
The highest PEEP, cmH ₂ O (median, (IQR))	5.8 (5.0-8.0)	5.9 (5.0-8.0)	0.74
The highest tidal volume, mL (median, (IQR))	539.0 (480.0-620.0)	483.5 (428.2-558.0)	<0.01
The highest plat pressure, cmH ₂ O (median, (IQR))	20.0 (16.0-24.0)	18.0 (15.0-22.0)	<0.01
The highest minute volume, L/min (median, (IQR))	9.0 (8.1-9.7)	8.8 (7.8-9.6)	<0.01
The highest drive pressure, cmH ₂ O (median, (IQR))	13.0 (10.0-16.0)	11.5 (9.0-14.0)	<0.01
Severity score and additional hemodynamic support:			
Vasopressors (n(%))	1,957(49.1%)	585 (48.3%)	0.67
Clinical outcomes:			
Need for dialysis or RRT during ICU stay (n(%))	203 (5.1%)	104 (8.6%)	<0.01
Length of ICU stay, days (median, (IQR))	5.7 (3.0-10.6)	6.1 (1.3-20.2)	<0.05
Length of hospital stay, days	11.7 (6.9-19.0)	13.0 (2.4-35.8)	<0.01

(median, (IQR))			
ICU mortality rate(n(%))	701 (17.6%)	220 (18.2%)	0.67
Hospital mortality rate(n(%))	1,084 (27.2%)	293 (24.2%)	<0.05

Continuous variables were expressed as the median and interquartile range (IQR). Categorical variables are expressed as absolute (n) and relative (%) frequency.

Abbreviations:

AKI-1: AKI stage 1, AKI-2: AKI stage 2, AKI-3: AKI stage 3, AKI-23: AKI stages 2 or 3, AKI-123: AKI stages 1 or 2 or 3. AKI stages were based on serum creatinine criteria of KDIGO (Kidney Disease: Improving Global Outcomes).

RRT: renal replacement therapy

FIO₂: Fraction of inspiration O₂

PEEP: Positive end expiratory pressure

The calculation of ventilator-free days was done considering only ICU length of stay and hospital length of stay due to the data availability in the database.

Table 2: Performance of clinical prediction models in development cohorts for AKI-23 and AKI-3 based on logistic regression and random forest.

Model	AKI-23		Model	AKI-3	
Machine learning algorithms	Logistic regression	Random forest	Machine learning algorithms	Logistic regression	Random forest
AUC	0.77 (0.69-0.84)	0.82 (0.76-0.88)	AUC	0.79 (0.70-0.88)	0.89 (0.84-0.95)
Sensitivity	0.74 (0.57-0.88)	0.57 (0.55-1.00)	Sensitivity	0.63 (0.50-0.96)	0.79 (0.67-1.00)
Specificity	0.74 (0.62-0.86)	0.90 (0.42-0.94)	Specificity	0.85 (0.53-0.91)	0.86 (0.62-0.92)
Positive predictive value	0.25 (0.20-0.36)	0.39 (0.17-0.53)	Positive predictive value	0.21 (0.11-0.32)	0.27 (0.14-0.40)
Negative predictive value	0.96 (0.94-0.98)	0.95 (0.94-1.00)	Negative predictive value	0.97 (0.97-1.00)	0.98 (0.98-1.00)
Classification Threshold (%)	10.2%	23.5%	Classification Threshold (%)	8.3%	11.9%

Classification Threshold based on the maximum Youden index. The 95% corresponding bootstrap confidence intervals (CIs) were computed. AKI stages were based on serum creatinine criteria of KDIGO (Kidney Disease: Improving Global Outcomes). AKI-23: AKI stages 2 or 3. AKI-3: AKI stages 3.

Abbreviations: AKI, acute kidney injury; AUC, area under the ROC curve; ROC, receiver operator characteristic.

Table 3: Performance of clinical prediction models in the internal validation cohorts for AKI-23 and AKI-3 based on random forest.

Model	AKI-23	Model	AKI-3
AUC	0.78 (0.74-0.82)	AUC	0.81 (0.76-0.87)
Sensitivity	0.47 (0.39-0.55)	Sensitivity	0.65 (0.54-0.75)
Specificity	0.89 (0.87-0.91)	Specificity	0.86 (0.84-0.88)
Positive predictive value	0.39 (0.34-0.45)	Positive predictive value	0.25 (0.21-0.29)
Negative predictive value	0.92 (0.91-0.93)	Negative predictive value	0.97 (0.96-0.98)
Classification Threshold (%)	23.5%	Classification Threshold (%)	11.9%

Classification Threshold based on the maximum Youden index in the development cohort. The 95% corresponding bootstrap confidence intervals (CIs) were computed. AKI stages were based on serum creatinine criteria of KDIGO (Kidney Disease: Improving Global Outcomes). AKI-23: AKI stages 2 or 3. AKI-3: AKI stages 3.

Abbreviations: AKI, acute kidney injury; AUC, area under the ROC curve; ROC, receiver operator characteristic.

Table 4 Performance of clinical prediction models in the external validation cohorts for AKI-23 and AKI-3 based on random forest.

Model	AKI-23	Model	AKI-3
AUC	0.80 (0.76-0.84)	AUC	0.80 (0.73-0.86)
Sensitivity	0.57 (0.48-0.66)	Sensitivity	0.67 (0.55-0.78)
Specificity	0.83 (0.81-0.85)	Specificity	0.85 (0.83-0.87)
Positive predictive value	0.23 (0.20-0.27)	Positive predictive value	0.18 (0.15-0.21)
Negative predictive value	0.96 (0.95-0.97)	Negative predictive value	0.98 (0.98-0.99)
Classification Threshold (%)	23.5%	Classification Threshold (%)	11.9%

Classification Threshold based on the maximum Youden index in the development cohort. The 95% corresponding bootstrap confidence intervals (CIs) were computed. AKI stages were based on serum creatinine criteria of KDIGO (Kidney Disease: Improving Global Outcomes). AKI-23: AKI stages 2 or 3. AKI-3: AKI stages 3.

Abbreviations: AKI, acute kidney injury; AUC, area under the ROC curve; ROC, receiver operator characteristic.

Figures

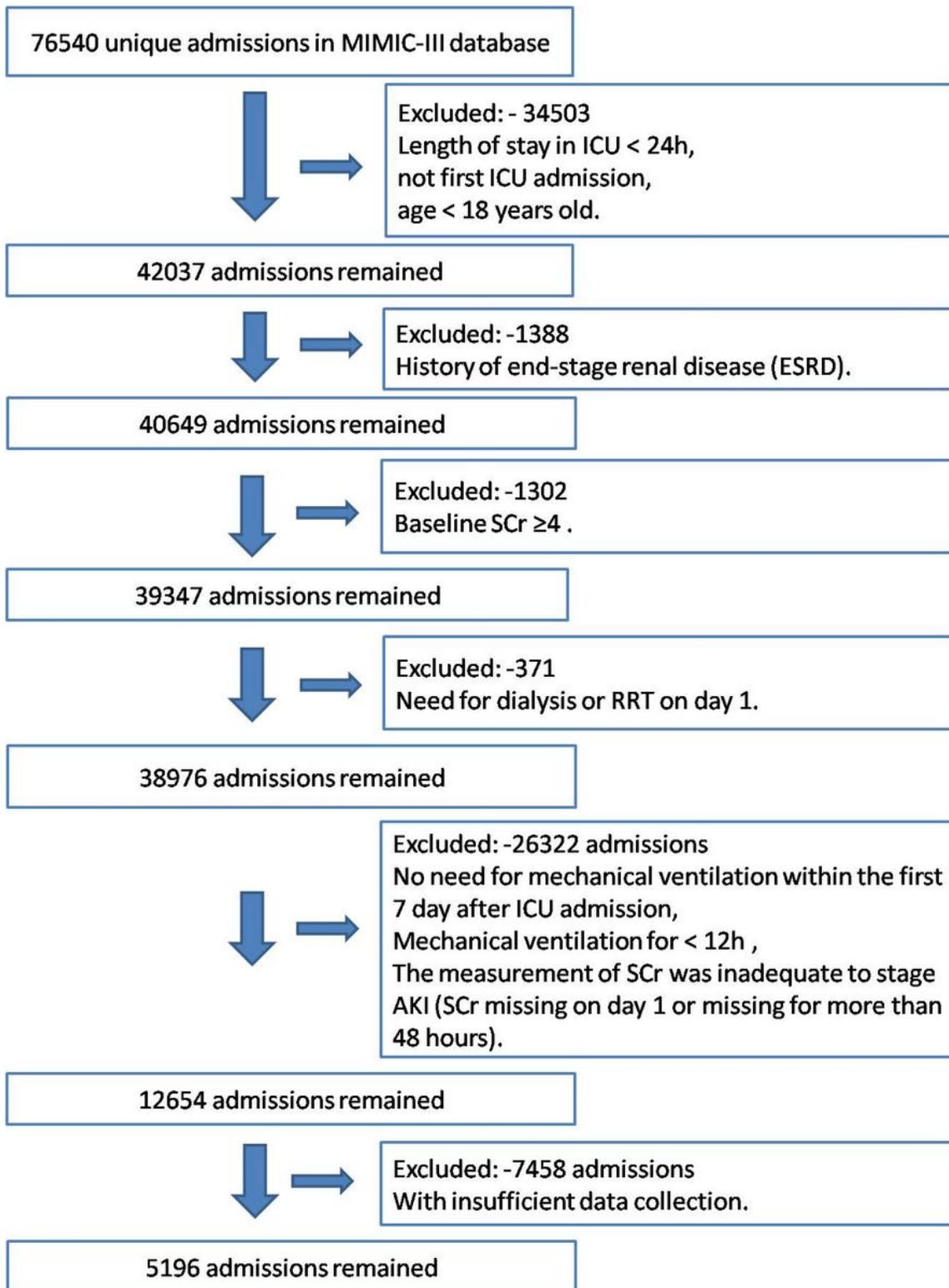


Figure 1

Study cohort selection workflow of MIMIC-IV database based on the designed inclusion and exclusion criteria. AKI: acute kidney injury, SCr: serum creatinine.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplemental.pdf](#)