Comparison of Approaches for Prediction of Renal Replacement Therapy-Free Survival in Patients with Acute Kidney Injury

Pattharawin Pattharanitima^{a, b} Akhil Vaid^a Suraj K. Jaladanki^c Ishan Paranjpe^c
Ross O'Hagan^c Kinsuk Chauhan^a Tielman T. Van Vleck^d Aine Duffy^d
Kumardeep Chaudhary^d Benjamin S. Glicksberg^e
Javier A. Neyra^f Steven G. Coca^a Lili Chan^a Girish N. Nadkarni^{a, d}

^aDepartment of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^bDepartment of Internal Medicine, Faculty of Medicine, Thammasat University, Pathum Thani, Thailand; ^cIcahn School of Medicine at Mount Sinai, New York, NY, USA; ^dDepartment of Genetics and Genomic Sciences, Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^eDepartment of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^fDepartment of Internal Medicine, University of Kentucky College of Medicine, Lexington, KY, USA

Keywords

Machine learning · Continuous renal replacement therapy · Discontinuation · Mortality · Acute kidney injury

Abstract

Background/Aims: Acute kidney injury (AKI) in critically ill patients is common, and continuous renal replacement therapy (CRRT) is a preferred mode of renal replacement therapy (RRT) in hemodynamically unstable patients. Prediction of clinical outcomes in patients on CRRT is challenging. We utilized several approaches to predict RRT-free survival (RRTFS) in critically ill patients with AKI requiring CRRT. **Methods:** We used the Medical Information Mart for Intensive Care (MIMIC-III) database to identify patients ≥ 18 years old with AKI on CRRT, after excluding patients who had ESRD on chronic dialysis, and kidney transplantation. We defined RRTFS as patients who were discharged alive and did not require RRT ≥ 7 days prior to hospital discharge. We utilized all available biomedical data up to CRRT initiation. We evalu-

ated 7 approaches, including logistic regression (LR), random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), multilayer perceptron (MLP), and MLP with long short-term memory (MLP + LSTM). We evaluated model performance by using area under the receiver operating characteristic (AUROC) curves. **Results:** Out of 684 patients with AKI on CRRT, 205 (30%) patients had RRTFS. The median age of patients was 63 years and their median Simplified Acute Physiology Score (SAPS) II was 67 (interquartile range 52–84). The MLP + LSTM showed the highest AUROC (95% CI) of 0.70 (0.67–0.73), followed by MLP 0.59 (0.54–0.64), LR 0.57 (0.52–0.62), SVM 0.51 (0.46–0.56), AdaBoost 0.51 (0.46–0.55), RF 0.44 (0.39–0.48), and XGBoost 0.43 (CI 0.38–0.47). **Conclusions:** A MLP + LSTM model outperformed other approaches for predicting RRTFS. Performance could be further improved by incorporating other data types.

© 2021 S. Karger AG, Basel

L.C. and G.N.N. contributed equally.

Introduction

Acute kidney injury (AKI) is a common complication in critically ill patients affecting up to 67% of patients and is associated with increased mortality [1]. Continuous renal replacement therapy (CRRT) is the preferential mode of renal replacement therapy (RRT) in critically ill patients who are hemodynamically unstable [2]. CRRT requires specialized equipment and personnel for intensive monitoring which generally results in greater in-hospital resource consumption than intermittent hemodialysis [3]. Unfortunately, mortality remains high in patients who require CRRT, reported up to 63% [4]. While there is not a widely accepted tool for prediction of adverse events in patients initiated on CRRT, the utility of such predictive approaches is extensive. Moreover, given that 25% of patients initiated on CRRT remain dialysis dependent on hospital discharge, identification of those who are likely to need chronic dialysis may aid patients and families in their care decisions [5, 6].

Disease severity scores that have been commonly used for mortality prediction in critically ill patients, such as Acute Physiology and Chronic Health Enquiry II (APACHE II), Sequential Organ Failure Assessment (SOFA), and Simplified Acute Physiology Score (SAPS), have poor performance in patients receiving CRRT [7, 8]. Previous studies in this population have shown that various parameters including effluent volume, fluid overload, net ultrafiltration rate, CRRT duration, urine output, serum creatinine, and biomarkers are associated with mortality and/or RRT-free survival (RRTFS); however, results were inconsistent. In addition, most of these studies used data obtained during or after CRRT initiation which is not practical for decisions that need to be made before CRRT initiation [9–15].

Machine learning algorithms detect underlying patterns in the data presented to them, and their performance at such tasks improves with exposure to more data [16]. However, there are limited studies in the field of kidney disease utilizing machine learning models [17, 18]. We tested the performance of standard logistic modeling and supervised machine learning algorithms to predict RRTFS after CRRT initiation in critically ill patients with AKI.

Materials and Methods

Data Source

We used publicly available data from the Medical Information Mart for Intensive Care III (MIMIC-III) database comprising of 61,532 intensive care unit (ICU) admissions from 46,520 patients

from the Beth Israel Deaconess Medical Center in Boston, MA, between 2001 and 2012 [19]. The database includes patient demographics, hourly vital signs, clinical measurements, laboratory results, treatments, diagnosis and procedural codes. Since the MIMIC-III database is publicly available and contains de-identified data, the paper is exempt from Ethics Committee review.

Study Population

We included patients who were ≥ 18 years old with AKI requiring CRRT. We excluded patients who only received slow continuous ultrafiltration and patients with International Classification of Diseases, Ninth revision diagnostic (ICD-9) codes of ESRD on chronic dialysis or kidney transplant. We identified patients with AKI by presence of ICD-9 codes of 584, 584.5–584.9 (see details in online suppl. Table 1; for all online suppl. material, see www.karger.com/doi/10.1159/000513700). We followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) reporting guideline [20].

Primary Outcome

The primary outcome was RRTFS, defined as patients who were discharged alive and did not require any RRT in the form of hemodialysis, peritoneal dialysis, or CRRT for 7 days or more prior to hospital discharge. Example cases in defining RRTFS are shown in online suppl. Fig. 1.

Data and Machine Learning Models

We used 7 different machine learning methods including logistic regression (LR), random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), multilayer perceptron (MLP), and MLP with long short-term memory (MLP + LSTM). LR uses the logistic function to classify data into 2 groups. RF utilizes random subsets and multiple features to construct a multitude of decision trees to classify data [21]. SVM constructs a decision boundary out of a hyperplane that allows for maximal separation of datapoints on either side of that boundary [22]. AdaBoost uses a similar principle of decision trees and enhances prediction accuracy by assigning different weights to each weak classifier and combining sets of weak classifiers into a single strong classifier [23]. XGBoost is comparable to AdaBoost; however, its algorithm is parallelizable and distributable which allows the model to work faster and be more suited for a large dataset [24]. MLP is a deep learning model. It consists of an input layer, one or more hidden layers, and an output layer. Each layer contains multiple nodes that are connected to the nodes on the next layer which allow data to flow from input layer to output layer. These layers and connections are the core structure for calculation of MLP [25]. The MLP + LSTM consists of 2 components that incorporate categorical data into the MLP segment and time series data into the LSTM segment. The LSTM incorporates time-dependent data through a form of recurrent neural networks, called LSTM units, which use a memory cell inside each neuron to remember prior inputs and regulate information input and output [25].

Features obtained from hospital admission and up to CRRT initiation were included for all models (online suppl. Table 2). Static features, such as demographic data, comorbidities, and initial CRRT mode were included. For all models except MLP + LSTM, time series data such as laboratory values, vital signs, and urine outputs from ICU admission and up to CRRT initiation were included as median values into the model. Continuous missing data

Table 1. The characteristics of adult patients with AKI who underwent CRRT in the study

	Patients with RRTFS (<i>n</i> = 205)	Patients without RRTFS (<i>n</i> = 479)
Male, <i>n</i> (%)	129 (63)	314 (66)
Age, years, median (IQR)	60.4 (49–71)	64.3 (53–76)
Weight, kg, median (IQR)	91.6 (78.5–104.9)	93.3 (80.4–114.4)
Ethnicity, <i>n</i> (%)		
White	154 (81)	313 (65)
Black/African American	18 (9)	31 (6)
Hispanic/Latino	8 (3)	11 (3)
Other	33 (11)	124 (26)
Charlson Comorbidity Index, median (IQR)	6 (4–7)	5 (3–7)
SAPS II, median (IQR)	63 (48–78)	69 (53–85)
Comorbidities, <i>n</i> (%)		
Hypertension	165 (80)	421 (88)
Diabetes (required insulin treatment)	73 (36)	158 (33)
Diabetes (required oral agent treatment)	63 (31)	101 (21)
Coronary artery disease	64 (31)	221 (43)
Congestive heart failure	58 (28)	126 (26)
Arrhythmias	51 (25)	121 (25)
Myocardial infarction	34 (17)	86 (18)
Anemia	48 (23)	105 (22)
Chronic obstructive pulmonary disease	27 (13)	119 (25)
Alcoholism	78 (38)	127 (27)
Liver failure	35 (17)	118 (25)
Smoker	28 (14)	80 (17)
Gastrointestinal bleeding	41 (20)	53 (11)
Hepatitis	36 (18)	85 (18)
Asthma	35 (17)	77 (17)
Renal failure	41 (20)	78 (16)
Mode of first CRRT session, <i>n</i> (%)		
CVVH	3 (1)	14 (3)
CVVHD	6 (3)	6 (1)
CVVHDF	92 (45)	191 (40)
Unknown	101 (49)	268 (56)
Parameters prior to CRRT initiation, median (IQR)		
Ventilator duration, day	9 (7–13)	10 (6–17)
Glasgow Coma Scale	10 (6–14)	10 (6–11)
SBP, mm Hg	101 (76–106)	127 (106–135)
DBP, mm Hg	50 (41–57)	57 (46–69)
BUN, mg/dL	47 (29–71)	52 (31–81)
Cr, mg/dL	2.6 (1.7–3.9)	2.4 (1.5–3.6)
Hb, g/dL	9.4 (8.5–10.3)	9.6 (8.8–10.6)
Lactate, mmol/L	2.0 (1.4–3.3)	2.7 (1.7–5.6)
Parameters at CRRT initiation		
Required mechanical ventilator	146 (71)	371 (77)
Required any vasopressor, <i>n</i> (%)	164 (80)	390 (81)
Loop diuretic, mg/day, median (IQR)*	56 (17–100)	40 (20–98)
In-hospital mortality, <i>n</i> (%)	0	384 (80)

AKI, acute kidney injury; BW, body weight; IQR, interquartile range; SAPS, Simplified Acute Physiology Score; CRRT, continuous renal replacement therapy; CVVH, continuous veno-venous hemofiltration; CVVHD, continuous veno-venous hemodialysis; CVVHDF, continuous veno-venous hemodiafiltration; Hb, hemoglobin; MAP, mean arterial blood pressure; RRTFS, renal replacement therapy-free survival. * Equivalent dose to intravenous furosemide.

Table 2. The performances of machine learning models for predicting RRTFS

	AUROC (95% CI)	Sensitivity			Specificity			PPV			NPV		
		20% cutoff	30% cutoff	50% cutoff	20% cutoff	30% cutoff	50% cutoff	20% cutoff	30% cutoff	50% cutoff	20% cutoff	30% cutoff	50% cutoff
LR	0.57 (0.52–0.62)	1.00	0.51	0.00	0.00	0.59	1.00	0.30	0.35	–	–	0.74	0.70
XGBoost	0.43 (0.38–0.47)	0.78	0.38	0.13	0.20	0.50	0.86	0.29	0.24	0.28	0.67	0.65	0.70
RF	0.44 (0.39–0.48)	0.69	0.40	0.17	0.24	0.51	0.82	0.28	0.26	0.29	0.64	0.67	0.70
AdaBoost	0.51 (0.46–0.55)	1.00	1.00	0.99	0.00	0.00	0.01	0.30	0.30	0.30	–	–	0.71
SVM	0.51 (0.46–0.56)	0.98	0.41	0.01	0.01	0.56	0.99	0.30	0.29	0.33	0.57	0.69	0.70
MLP	0.59 (0.54–0.64)	1.00	0.62	0.00	0.00	0.57	1.00	0.31	0.42	–	–	0.81	0.69
MLP + LSTM	0.70 (0.67–0.73)	1.00	0.71	0.00	0.00	0.59	1.00	0.31	0.42	–	–	0.81	0.69

RRTFS, renal replacement therapy-free survival; AdaBoost, adaptive boosting; AUROC, area under the receiver operating characteristics; CI, confidence interval; RF, random forest; MLP + LSTM, multilayer perceptron with long short-term memory; LR, logistic regression; MLP, multilayer perceptron; NPV, negative predictive value; PPV, positive predictive value; RF, random forest; SVM, support vector machine; XGBoost, extreme gradient boosting.

were replaced with K nearest neighbor imputation with $K = 5$. Data normalization was performed using Z-score standardization.

For MLP + LSTM and MLP, static features were fed into MLP with 2 fully connected layers of 100 nodes and 50 nodes, respectively. The results were then concatenated and passed through a fully connected layer of 10 nodes and then to the output layer. All fully connected layers used rectified linear unit activation and the output layer used sigmoid activation. The time series data included mean, standard deviation, count, minimum and maximum of laboratory values, vital signs, drug input events, and urine outputs over 6-h intervals which were fed into the LSTM with each 100-cell layer in series and a fully connected layer of 50 cells. Missing data were replaced with zero and no imputation was performed. The model was trained for 250 epochs with early stopping enabled with a batch size of 16 and a validation split of 0.1. No feature selection was performed.

Statistical Analysis

Baseline data were presented as number (%) and median (interquartile range [IQR]) for categorical and continuous variables, respectively. Model performances were evaluated by calculating area under the receiver operating characteristic (AUROC) curve with 95% confidence interval (95% CI), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). We performed k -fold cross-validation procedure ($k = 25$) to calculate the average AUROC and 95% CI for each model. Since there is no standard prognostic score for RRTFS, we used LR as a baseline comparison. The database queries and the analyses were performed using PostgreSQL software version 11.3 (PostgreSQL Global Development Group, Berkeley, CA, USA), Python software version 3.6.8 (Python Software Foundation, DE, USA), and Stata/IC 15.1 software (StataCorp, College Station, TX, USA).

Results

Study Population Characteristics

As shown in online suppl. Fig. 2, 684 out of 806 patients were included, and 115 and 7 patients were excluded

as they were ESRD on chronic dialysis or received slow continuous ultrafiltration, respectively. The median (IQR) age of the patients was 63 (52–75) years; 443 (65%) were male, and 467 (68%) were white. The most common initial CRRT mode was continuous veno-venous hemodiafiltration ($N = 283$ [41%]). The median (IQR) SAPS II was 67 (52–84). Vasopressors were used in 554 (81%) patients. There were 205 (30%) patients who had RRTFS. Among 479 patients without RRTFS, 384 (80%) patients had in-hospital mortality. Patient characteristics are shown in Table 1.

Model Prediction of RRTFS

A total of 81 features from hospital admission and up to CRRT initiation were included into machine learning models (online suppl. Table 2). The AUROCs and 95% CIs are shown in Table 2. Since there are no universally accepted model for RRTFS prediction, the LR model served as a baseline for model comparison and was found to have an AUROC of 0.57 (95% CI 0.52–0.62). Among the 6 other machine learning models, only MLP + LSTM and MLP had a higher AUROC than LR. The MLP + LSTM gave the best discrimination with an AUROC of 0.70 (95% CI 0.67–0.73), followed by MLP 0.59 (95% CI 0.54–0.64), SVM 0.51 (95% CI 0.46–0.56), AdaBoost 0.51 (95% CI 0.46–0.55), RF 0.44 (95% CI 0.39–0.48), and XGBoost 0.43 (95% CI 0.38–0.47).

The sensitivity, specificity, PPV, and NPV were shown according to predicted risk thresholds of 20, 30, and 50% (Table 2). The MLP + LSTM gave the highest sensitivity for 20% cutoff (1.00) and second highest sensitivity for 30% cutoff (0.71). It gave the highest specificity for 30 and 50% cutoff (0.59 and 1.00). It gave the highest PPV

for 20 and 30% cutoffs (0.31 and 0.42), highest NPV for 30% cutoff (0.81), and third highest NPV for 50% cutoff (0.69).

Discussions

This study demonstrates the performances of 7 approaches in predicting RRTFS in patients with AKI requiring CRRT. We explored several demographic and clinical features including various time series data. Our results showed that deep learning models surpassed other models, with AUROC improving by 13% for MLP + LSTM compared to LR.

Among machine learning models, MLP + LSTM is the only model that can aggregate time series data over 6-h intervals from hospital admission to the time of CRRT initiation and incorporate them into the model chronologically, while other models only included the median of each time series data. Singh et al. [26] showed that incorporating time series data from patient medical history could improve model performance in predicting kidney function deterioration in diabetic patients. These results correspond to ours, where the MLP + LSTM outperformed other machine learning models. This finding suggests that understanding changes in features over time are important for accurate outcome prediction.

While currently used severity illness scoring systems, such as APACHE II, SOFA, and SAPS are practical since they utilize variables obtained early in ICU admission, their performances are poor in patients receiving CRRT [7]. From our dataset, the median SAPS II was 67 which corresponds to an in-hospital mortality rate of 80% [27], whereas the actual mortality rate was only 56%. These discordances were also found in previous studies conducted in CRRT populations [10, 11, 28]. Thus, these scoring systems are not compatible for CRRT patients and development of a prediction model is necessary to guide clinical care and decision-making.

Kang et al. [29] used machine learning to predict mortality in patients initiated on CRRT. They included clinical data at the time of CRRT initiation into models and compared their performance with standard severity illness scoring systems. They demonstrated that all machine learning models achieved better results than traditional severity illness scoring systems, particularly the artificial neural network (ANN) model which outperformed other machine learning models. While most research has focused on mortality, we chose our outcome of RRTFS which is more informative to patients and fam-

ily members as the need for chronic dialysis has a significant impact on quality of life. Unlike our study, Kang et al. [29] found that RF performed the best followed by ANN, whereas the MLP + LSTM and MLP which contain ANN structures were the 2 best performing models in our study. Although there were methodological differences between their study and ours, these findings emphasize the valuable of machine learning in predicting outcomes in healthcare settings, particularly in patients receiving CRRT.

Due to the near universal use of electronic health records (EHR), unprecedented amounts of biomedical data are now available for predictive modeling. However, conventional statistical methods have several limitations for leveraging this data. First, conventional statistics usually assume data is normally distributed, independent, and linearly associated, which does not hold true in EHR data [30]. Second, they are model driven as opposed to data driven. While this is appropriate for hypothesis-driven research with a limited number of variables, EHR data allows for the assessment of many correlated variables and their combinations for modeling. Third, there are computational and theoretical limitations in dealing with datasets with large numbers of variables and subjects. Last, the variables must be preselected because of dimensionality issues. Therefore, there is a possibility of losing critical information relevant to outcome predictions [22]. In the setting of patients with AKI requiring CRRT, a large amount of data is generated resulting in a multitude of variables that may have clinical significance.

However, our study should be considered in the light of a few limitations. First, the overall number of included patients in the present study is modest and training the models with a larger sample size dataset may improve accuracy of the results. Second, this study was done using single-center dataset which decreases the generalizability of our models. External validation is necessary before any potential clinical implementation. Third, we used a constant number of input features. Further feature selection to remove redundant features or add potentially relevant features such as microbiological or radiological results may improve model performance. Fourth, our method of outcome assessment will result in varying in duration of hospital stay and duration of RRT for each patient, and this could lead to bias. Finally, a predictive accuracy of 70% may not be sufficient for clinical implementation. However, the notable performance improvements of deep learning approaches suggest that these models may benefit from the addition of multi-modal data types in-

cluding physiological waveform data and data from the CRRT machine.

In conclusion, this study suggests that deep learning models, particularly MLP + LSTM, outperformed other approaches for the prediction of RRTFS in AKI patients requiring CRRT.

Statement of Ethics

The MIMIC-III database is publicly available and contains de-identified data, so the paper is exempt from Ethical Committee approval.

Conflict of Interest Statement

G.N.N. has received operational funding from Goldfinch Bio in the past 3 years. G.N.N. receives financial compensation as consultants and advisory board members for RenalytixAI and owns equity and stock options in RenalytixAI. G.N.N. is a scientific co-founder of RenalytixAI. S.G.C. reports consultancy agreements from RenalytixAI, Takeda Pharmaceuticals, CHF Solutions, Relypsa, Bayer, Boehringer Ingelheim, Quark, and Akebia. S.G.C. reports ownership interest from RenalytixAI and pulseData. S.G.C.

receives research funding from RenalytixAI and inRegen. S.G.C. reports patents and interventions from RenalytixAI. S.G.C. is a scientific advisor or membership for RenalytixAI. S.G.C. owns part of RenalytixAI. T.T.V.V. receives financial compensation as a consultant for Clinithink and owns equity in Clinithink. J.A.N. has received consulting fees from Baxter Healthcare in the past 3 years. Other authors declare that they have no conflict of interest.

Funding Sources

This project was supported in part by the NIH (R56DK126930).

Author Contributions

P.P., A.V., S.K.J., K.C., K.C., S.G.C., L.C., and G.N.N. were involved in data interpretation. P.P., A.V., S.K.J., I.P., R.O.H., T.T.V.V., A.D., S.G.C., L.C., and G.N.N. were involved in data analysis. P.P., A.V., S.K.J., S.G.C., L.C., and G.N.N. were involved in manuscript drafting. S.G.C., L.C., and G.N.N. were involved in supervision and mentorship. P.P., A.V., S.K.J., I.P., R.O.H., K.C., T.T.V.V., A.D., K.C., B.S.G., J.A.N., S.G.C., L.C., and G.N.N. were involved in study design and contributed important intellectual content during manuscript revision. All authors have given their approval to the final draft.

References

- 1 Hoste EA, Clermont G, Kersten A, Venkataraman R, Angus DC, De Bacquer D, et al. RIFLE criteria for acute kidney injury are associated with hospital mortality in critically ill patients: a cohort analysis. *Crit Care*. 2006; 10(3):R73.
- 2 Kidney Disease Improving Global Outcome (KDIGO) Acute Kidney Injury Working Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl*. 2012;2: 1–138.
- 3 Garay OU, Palacios A, Pichon-Riviere A, Augustovski F, Martí SG, Hernández-Vásquez A, et al. The cost-effectiveness of continuous versus intermittent renal replacement therapies in acute kidney injury: perspective of the social services for the elderly in Argentina. *Value Health Reg Issues*. 2019 Dec;20:142–8.
- 4 Prasad B, Urbanski M, Ferguson TW, Karremann E, Tangri N. Early mortality on continuous renal replacement therapy (CRRT): the prairie CRRT study. *Can J Kidney Health Dis*. 2016;3:36.
- 5 Schneider AG, Bellomo R, Bagshaw SM, Glassford NJ, Lo S, Jun M, et al. Choice of renal replacement therapy modality and dialysis dependence after acute kidney injury: a systematic review and meta-analysis. *Intensive Care Med*. 2013 Jun;39(6):987–97.
- 6 de Souza Oliveira MA, Dos Santos TOC, Monte JCM, Batista MC, Pereira VG, Dos Santos BFC, et al. The impact of continuous renal replacement therapy on renal outcomes in dialysis-requiring acute kidney injury may be related to the baseline kidney function. *BMC Nephrol*. 2017 May;18(1):150.
- 7 da Hora Passos R, Ramos JG, Mendonça EJ, Miranda EA, Dutra FR, Coelho MF, et al. A clinical score to predict mortality in septic acute kidney injury patients requiring continuous renal replacement therapy: the HELE-NICC score. *BMC Anesthesiol*. 2017 Feb; 17(1):21.
- 8 Kim Y, Park N, Kim J, Kim DK, Chin HJ, Na KY, et al. Development of a new mortality scoring system for acute kidney injury with continuous renal replacement therapy. *Nephrology*. 2019 Dec;24(12):1233–40.
- 9 Fayad AI, Buamscha DG, Ciapponi A. Intensity of continuous renal replacement therapy for acute kidney injury. *Cochrane Database Syst Rev*. 2016 Oct;10:CD010613.
- 10 Gaudry S, Hajage D, Schortgen F, Martin-Lefevre L, Pons B, Boulet E, et al. Initiation strategies for renal-replacement therapy in the intensive care unit. *N Engl J Med*. 2016 Jul; 375(2):122–33.
- 11 Zarbock A, Kellum JA, Schmidt C, Van Aken H, Wempe C, Pavenstädt H, et al. Effect of early vs delayed initiation of renal replacement therapy on mortality in critically ill patients with acute kidney injury: the ELAIN randomized clinical trial. *JAMA*. 2016 May; 315(20):2190–9.
- 12 Chen X, Chen Z, Wei T, Li P, Zhang L, Fu P. The effect of serum neutrophil gelatinase-associated lipocalin on the discontinuation of continuous renal replacement therapy in critically ill patients with acute kidney injury. *Blood Purif*. 2019 Mar;48(1): 10–7.
- 13 Katayama S, Uchino S, Uji M, Ohnuma T, Namba Y, Kawarazaki H, et al. Factors predicting successful discontinuation of continuous renal replacement therapy. *Anaesth Intensive Care*. 2016 Jul;44(4):453–7.
- 14 Woodward CW, Lambert J, Ortiz-Soriano V, Li Y, Ruiz-Conejo M, Bissell BD, et al. Fluid overload associates with major adverse kidney events in critically ill patients with acute kidney injury requiring continuous renal replacement therapy. *Crit Care Med*. 2019; 47(9):e753–60.
- 15 Murugan R, Kerti SJ, Chang CH, Gallagher M, Clermont G, Palevsky PM, et al. Association of net ultrafiltration rate with mortality among critically ill adults with acute kidney injury receiving continuous venovenous hemodiafiltration: a secondary analysis of the randomized evaluation of normal vs augmented level (RENAL) of renal replacement therapy trial. *JAMA Netw Open*. 2019 Jun 5; 2(6):e195418.
- 16 Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018; 15(4):233–4.

- 17 Nadkarni GN, Chaudhary K, Coca SG. Machine learning in glomerular diseases: promise for precision medicine. *Am J Kidney Dis*. 2019 Sep;74(3):290–2.
- 18 Saez-Rodriguez J, Rinschen MM, Floege J, Kramann R. Big science and big data in nephrology. *Kidney Int*. 2019 Jun;95(6):1326–37.
- 19 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 May;3:160035.
- 20 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015 Jan;350:g7594.
- 21 Breiman L. Random forests. *Mach Learn*. 2001 Oct;45(1):5–32.
- 22 Shouval R, Bondi O, Mishan H, Shimoni A, Unger R, Nagler A. Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT. *Bone Marrow Transplant*. 2014 Mar;49(3):332–7.
- 23 Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. *Ann Stat*. 2000;28(2):337–407.
- 24 Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *CoRR*. 2016:785–94.
- 25 Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinformatics*. 2017 Sep 1; 18(5):851–69.
- 26 Singh A, Nadkarni G, Gottesman O, Ellis SB, Bottinger EP, Guttag JV. Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration. *J Biomed Inform*. 2015 Feb;53:220–8.
- 27 Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993 Dec;270(24):2957–63.
- 28 Palevsky PM, Palevsky PM, Zhang JH, O'Connor TZ, Chertow GM, Crowley ST, et al. Intensity of renal support in critically ill patients with acute kidney injury. *N Engl J Med*. 2008 Jul;359(1):7–20.
- 29 Kang MW, Kim J, Kim DK, Oh KH, Joo KW, Kim YS, et al. Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy. *Crit Care*. 2020 Feb;24(1):42.
- 30 Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Statist Sci*. 2001 Aug;16(3):199–231.