Journal of
Health & Medical Informatics
ISSN: 2157-7420

Kiljanek et al., J Health Med Informat 2019, 10:5

Research Article Open Access

Prediction of Incident Renal Replacement Therapy (iRRT) Use in ICU Population with Artificial Intelligence (AI), and SOFA, OASIS, APSIII Severity Scores

Lukasz R Kiljanek*, Medstar Shah and Sandeep Aggarwal

Drexel University College of Medicine, Philadelphia, PA, USA

Abstract

Background: Prediction tool for incident renal replacement therapy (iRRT) use could potentially improve outcomes in ICU population. We used the data from the Medical Information Mart for Intensive Care III (MIMIC III) database to create artificial intelligence (AI) iRRT use prediction model.

Methods: Based on routinely collected data in ICU we identified and engineered 679 candidate predictors of iRRT use. The iRRT was defined as any dialysis-related event charted in the electronic medical record (EMR) within the seven days following the first 24 hours of ICU admission. ICU stays of patients on dialysis before ICU admission, and with dialysis-related events charted before the end of first 24 hours were excluded. Remaining 18379 ICU stays were randomly divided 400 times, into training and testing datasets. For each random training dataset, AI-model for iRRT prediction was trained. Predictions of AI, SOFA, OASIS, and APSIII, were validated on testing dataset against the known use of iRRT with the area under the curve (AUC) of receiver-operator characteristics curve (ROC) recorded.

Result: For all 400 iterations, AUC of ROC for AI-model was 0.88 [95% CI 0.88-0.89] and was higher than SOFA, APSIII and OASIS: 0.82 [0.82-0.82], 0.81 [0.81-0.81], and 0.7 [0.69-0.7] AUCs respectively ($p<0.001$).

Conclusion: AI-model was accurate in predicting patients who survive until, consent and undergo iRRT after ICU admission. High AUC for the AI model trained only on data from first 24 hours of ICU stay emphasizes the importance of initial ICU management on renal outcomes.

Keywords: Dialysis; Machine learning; Renal replacement therapy; Random forest; Outcomes; Artificial intelligence; Intensive care unit

Introduction

Acute kidney injury (AKI) commonly complicates ICU stays [1] and is associated with significant mortality [2] and commonly leads to iRRT use. Prediction tool for iRRT use, and not only iRRT being indicated, could potentially optimize the timing of care delivery, staffing planning, and allow for preemptive inter-hospital transfers. To the best of authors' knowledge, there is no functional, clinically available tool to predict iRRT in ICU population. Computer-based machine learning (ML) and AI models have been successfully employed to predict risk in medicine [3], and business [4]. We attempted to use random forest (RF), an AI-model, to create a prediction tool for iRRT use within the seven days following the first 24 hours of ICU stays. We employed data from Medical Information Mart for Intensive Care III (MIMIC 3) database [5].

Methods

Following applicable institutional IRB policies, we obtained access to the ICU MIMIC 3 [5], a relational database containing a wide range of de-identified clinical data on over 50,000 ICU stays at Beth Israel Deaconess Medical Center of Harvard Medical, Boston, MA. The database includes demographic, vital signs, laboratory values, medication administration records, notes, insurance information and other data. It has been successfully used for research [6] including AI modeling [7]. Data has been collected by two different recording systems: CareVue and Metavision. To simplify research design, we only used data collected by Metavision, which was used to record more recent portion of the database from 2008-2014. We included at least 24 hours long ICU stays for patients 16 years old and older. ICU stays of patients already on renal replacement therapy (RRT) and with dialysis-related events charted before the end of 24 hours were excluded. Final database consisted of 18379 ICU stays. For database management, we used pgAdmin4 [8] with PostgreSQL. Raw variables were uploaded into R-project [9] for further (AI) analyses. We used the following R project libraries: doSNOW [10], progress [11], tableone [12] and data.table [13].

Extracting raw variables from the database

During the design phase, the investigators have chosen clinical data points both firmly and rather liberally associated with the iRRT, but all routinely collected on ICU encounters. Examples include demographics, laboratory values, urine output, vital signs, weight, past medical history, admission type (elective vs. nonelective), use of mechanical ventilation, administration records of medications known to affect renal function. Supplementary Table 1 lists all variables. One of the most significant advantages of AI-modeling is that AI can automatically and efficiently handle a large number of variables, pick the most important and ignore less or not relevant. For continuous variables, visual inspection of histograms and clinical experience were used to exclude data points which were much more to be an error vs.

***Corresponding author:** Lukasz R Kiljanek, Drexel University College of Medicine, 3900 Ford Road, Apt. 11H. Philadelphia, PA 19131, USA, Tel: +1 484 885 6833; E-mail: lukaszkiljanek@gmail.com

Received November 13, 2019; **Accepted** November 29 2019; **Published** December 10, 2019

Citation: Kiljanek LR, Shah M, Aggarwal S (2019) Prediction of Incident Renal Replacement Therapy (iRRT) Use in ICU Population with Artificial Intelligence (AI), and SOFA, OASIS, APSIII Severity Scores. J Health Med Informat 10: 339.

Copyright: © 2019 Kiljanek LR, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

genuinely outlying measured value (for example one can safely assume that blood pH of 1 is an erroneous data point). Supplementary Table 1 list shows all cut-offs.

Engineered features: Mathematical and clinical

For each laboratory value and vital sign, additional mathematical features were engineered, as listed in Supplementary Table 2 (rows 1 through 10). Of note "*MaxSubMinPerMaxVariable=(MaxVariable–MinVariable/MaxVariable)*" was engineered to reflect the relational variability of a given variable. "*WtdMeanVariable*" is as a time-averaged variable - integral of levels divided by the 24 hours of time. DeltaVariable reflects the trend for the given variable within the first 24 hours. To further improve the accuracy of predictions (AUC), investigators engineered additional clinical features, reflecting clinical reasoning. Supplementary Table 2 (rows below 10th) lists full list of clinical variables. Table 1 contains most relevant, mathematical and clinical features.

Study population characteristics

Table 1 shows study population, stratified by the use of iRRT outcome. Chi-squared test for categorical variables (with continuity correction) and equal means in a one-way layout test for continuous variables (with equal variance assumption) derived the p-values [14].

	No iRRT	iRRT	p-value
N	18038	341	
(%) for categoric and mean (sd) for continuous variables			
Gender=M (%)	10145 (56.2)	194 (56.9)	0.854
Age (sd)	64.91 (17.14)	63.72 (15.38)	0.204
Ethnicity (%)			0.019
African American	1646 (9.1)	37 (10.9)	
Asian	482 (2.7)	8 (2.3)	
Caucasian	13375 (74.1)	230 (67.4)	
Other/Unknown	2535 (14.1)	66 (19.4)	
Admission type (%)			0.173
Elective	2617 (14.5)	38 (11.1)	
Emergency	15231 (84.4)	298 (87.4)	
Urgent	190 (1.1)	5 (1.5)	
History of asthma (%)	395 (2.2)	9 (2.6)	0.708
History of COPD (%)	800 (4.4)	12 (3.5)	0.495
History of CAD or angina OR mi (%)	1175 (6.5)	30 (8.8)	0.115
History of MI (%)	489 (2.7)	10 (2.9)	0.935
History of CHF (%)	787 (4.4)	32 (9.4)	<0.001
History of PVD (%)	321 (1.8)	14 (4.1)	0.003
History of hypertension (%)	2342 (13.0)	56 (16.4)	0.074
History of anemia (%)	555 (3.1)	19 (5.6)	0.014
History of arrythmias (%)	771 (4.3)	24 (7.0)	0.019
History of pacemaker (%)	238 (1.3)	7 (2.1)	0.352
History of CVA (%)	297 (1.6)	11 (3.2)	0.042
History of seizures (%)	191 (1.1)	2 (0.6)	0.562
History of diabetes on oral medications (%)	575 (3.2)	19 (5.6)	0.021
History of diabetes on insulin (%)	740 (4.1)	36 (10.6)	<0.001
History of GI bleed (%)	416 (2.3)	11 (3.2)	0.35
History of hepatitis (%)	313 (1.7)	12 (3.5)	0.023
History of liver failure (%)	241 (1.3)	14 (4.1)	<0.001
History of pancreatitis (%)	101 (0.6)	4 (1.2)	0.26
History of renal failure (%)	349 (1.9)	24 (7.0)	<0.001
History of hemo-or peritoneal dialysis (%)	0 (0.0)	0 (0.0)	
History of smoking (%)	628 (3.5)	9 (2.6)	0.488
History of Etoh (%)	501 (2.8)	14 (4.1)	0.191
Urine output within first 24 hours (sd)	1992.55 (1229.31)	1148.70 (1153.18)	<0.001
Urine output within first 24 hours divided by weight (sd)	25.86 (17.19)	14.19 (15.05)	<0.001
Ventilation within first 24 hours (%)	5976 (33.1)	205 (60.1)	<0.001
SBP *TWA (sd)	118.86 (15.88)	115.16 (16.63)	<0.001
DBP *TWA (sd)	62.37 (10.73)	59.64 (11.02)	<0.001
MAP *TWA (sd)	77.56 (10.68)	75.16 (10.57)	<0.001
HR *TWA (sd)	85.47 (15.56)	89.00 (17.76)	<0.001
RR *TWA (sd)	18.92 (3.96)	19.86 (4.34)	<0.001
O2 saturation*TWA (sd)	97.05 (1.99)	96.81 (2.38)	0.03

Weight *TWA (sd)	81.88 (23.63)	85.26 (21.81)	0.023
Dopamine – total dose administered (sd)	12.15 (122.94)	93.90 (486.73)	<0.001
Dopamine – number of times dose recorded in EMR (sd)	0.19 (1.52)	0.77 (3.22)	<0.001
Norepinephrine – total dose administered (sd)	1.28 (7.24)	7.51 (21.31)	<0.001
Norepinephrine – number of times dose recorded in EMR (sd)	1.02 (3.82)	3.82 (7.85)	<0.001
Dobutamine – total dose administered (sd)	1.26 (34.31)	11.17 (136.05)	<0.001
Dobutamine – number of times dose recorded in EMR (sd)	0.02 (0.37)	0.13 (1.07)	<0.001
Epinephrine – total dose administered (sd)	81.33 (8983.35)	0.23 (1.55)	0.868
Epinephrine – number of times dose recorded in EMR (sd)	0.14 (1.09)	0.35 (1.62)	<0.001
Phenylephrine – total dose administered (sd)	11.18 (59.27)	30.77 (93.27)	<0.001
Phenylephrine – number of times dose recorded in EMR (sd)	1.50 (4.50)	2.85 (6.52)	<0.001
Vasopressin – total dose administered (sd)	0.31 (4.89)	1.27 (7.44)	<0.001
Vasopressin – number of times dose recorded in EMR (sd)	0.03 (0.34)	0.16 (0.70)	<0.001
SCr (serum creatinine) *TWA (sd)	1.15 (0.87)	2.78 (1.92)	<0.001
K (potassium) *TWA (sd)	4.13 (0.53)	4.40 (0.63)	<0.001
SCr=minimal from all values recorded within the first 24 hours (sd)	1.09 (0.82)	2.57 (1.89)	<0.001
SCr=maximal from all values recorded within the first 24 hours (sd)	1.21 (0.94)	2.96 (1.99)	<0.001
SCr=arithmetic mean from SCr all values recorded within the first 24 hours (sd)	1.15 (0.88)	2.77 (1.92)	<0.001
SCr=last SCr minus first recorded within the first 24 hours (sd)	-0.02 (0.32)	0.33 (0.51)	<0.001
SCr=last SCr minus first recorded within the first 24 hours or zero if only one sCr available (sd)	-0.01 (0.26)	0.28 (0.49)	<0.001
SCr=first value recorded within the first 24 hours (sd)	1.16 (0.91)	2.63 (1.89)	<0.001
SCr=last value recorded within the first 24 hours (sd)	1.15 (0.86)	2.91 (1.97)	<0.001
pH= (blood pH on blood gasses) TWA (sd)	7.38 (0.07)	7.34 (0.08)	<0.001
gasBicarb (bicarb recorded on blood gasses) *TWA (sd)	21.77 (5.90)	18.48 (2.97)	0.081
BUN (blood urea nitrogen) *TWA (sd)	24.02 (18.93)	48.19 (31.43)	<0.001
Bicarb *TWA (sd)	24.26 (4.40)	21.53 (5.25)	<0.001
Lactate *TWA (sd)	1.93 (1.18)	2.81 (2.14)	<0.001
Calcium *TWA (sd)	8.28 (0.76)	8.23 (1.12)	0.206
Phos=maximal from all values recorded within the first 24 hours (sd)	3.68 (1.16)	5.20 (1.74)	<0.001
Phos=arithmetic mean from all values recorded within the first 24 hours (sd)	3.45 (1.05)	4.78 (1.63)	<0.001
Phos=last value recorded within the first 24 hours (sd)	3.41 (1.08)	4.84 (1.71)	<0.001
Phos *TWA (sd)	3.45 (1.04)	4.79 (1.63)	<0.001
CK (creatine kinase) *TWA (sd)	742.85 (3920.06)	3492.30 (17270.77)	<0.001
PTH *TWA (sd)	110.88 (154.27)	69.18 (38.44)	0.706
PTH=first value recorded within the first 24 hours (sd)	110.85 (154.29)	60.50 (26.16)	0.649
PTH=last value recorded within the first 24 hours (sd)	110.90 (154.25)	76.50 (48.79)	0.756
Phos=maximal from all values recorded within the first 24 hours (sd)	3.68 (1.16)	5.20 (1.74)	<0.001
Phos=arithmetic mean from all values recorded within the first 24 hours (sd)	3.45 (1.05)	4.78 (1.63)	<0.001
Phos=last Phos minus first recorded within the first 24 hours (sd)	-0.15 (1.00)	0.17 (1.20)	<0.001
Phos=last Phos minus first recorded within the first 24 hours or zero if only one Phos available (sd)	-0.08 (0.74)	0.12 (1.03)	<0.001
Phos=first recorded within the first 24 hours (sd)	3.49 (1.15)	4.72 (1.72)	<0.001
Phos=last recorded within the first 24 hours (sd)	3.41 (1.08)	4.84 (1.71)	<0.001
Phos=number of recorded values within the first 24 hours (sd)	1.58 (1.11)	2.36 (1.46)	<0.001
PTH=minimal from all values recorded within the first 24 hours (sd)	110.85 (154.29)	60.50 (26.16)	0.649
PTH=maximal from all values recorded within the first 24 hours (sd)	110.90 (154.25)	76.50 (48.79)	0.756
PTH=arithmetic mean from all values recorded within the first 24 hours (sd)	110.87 (154.27)	68.50 (37.48)	0.702
PTH=maximal minus minimal divided by maximal value recorded within the first 24 hrs (sd)	0.00 (0.03)	0.14 (0.20)	<0.001
PTH=last PTH minus first recorded within the first 24 hours (sd)	1.00 (1.00)	32.00 (NA)	
PTH=last PTH minus first recorded within the first 24 hours or zero if only one PTH available (sd)	0.05 (0.29)	16.00 (22.63)	<0.001
Glucose *TWA (sd)	134.46 (47.42)	144.71 (55.83)	<0.001
Chloride *TWA (sd)	105.54 (5.78)	103.84 (6.44)	<0.001
pO2 (oxygen partial pressure on blood gasses) *TWA (sd)	137.79 (67.65)	121.86 (57.36)	<0.001
pCo2 (carbon dioxide partial pressure on blood gasses) *TWA (sd)	41.56 (9.87)	39.12 (8.56)	<0.001
NTproBNP *TWA (sd)	7947.62 (10801.84)	24996.92 (17899.41)	<0.001

TroponineT *TWA (sd)	0.64 (1.83)	0.99 (2.18)	0.033
TP (total protein) *TWA (sd)	5.55 (1.02)	5.20 (1.24)	0.298
CRP (C-reactive protein) *TWA (sd)	85.29 (81.78)	63.76 (82.16)	0.561
Na (sodium) *TWA (sd)	138.70 (4.51)	137.67 (5.39)	<0.001
Albumin *TWA (sd)	3.11 (0.66)	2.93 (0.70)	<0.001
BiliTotal (total bilirubin) *TWA (sd)	1.96 (4.23)	3.65 (6.66)	<0.001
Mg (magnesium) *TWA (sd)	2.02 (0.34)	2.15 (0.44)	<0.001
PTT (activated partial thromboplastin time) *TWA (sd)	36.68 (18.11)	45.19 (22.79)	<0.001
INR (international normalized ration) *TWA (sd)	1.46 (0.78)	1.77 (0.95)	<0.001
ALP (alkaline phosphatase) *TWA (sd)	117.78 (136.62)	115.81 (91.58)	0.827
AST (aspartate aminotransferase) *TWA (sd)	175.08 (672.13)	557.94 (1620.54)	<0.001
ALT (alanine aminotransferase) *TWA (sd)	133.78 (544.73)	313.79 (914.48)	<0.001
CKMB (creatine kinase MB isoenzyme) *TWA (sd)	18.09 (43.58)	30.09 (72.18)	0.002
CKMB (creatine kinase MB isoenzyme index) *TWA (sd)	6.21 (4.76)	6.33 (4.90)	0.832
Haptoglobin concentration *TWA (sd)	185.92 (121.59)	152.42 (108.04)	0.214
ESR (erythrocyte sedimentation rate) *TWA (sd)	51.09 (39.74)	29.67 (27.90)	0.191
Ammonia *TWA (sd)	62.08 (64.65)	73.58 (57.59)	0.513
HBA1C (glycosylated hemoglobin A1C) *TWA (sd)	6.97 (2.22)	6.36 (0.74)	0.268
HGB (hemoglobin concentration) *TWA (sd)	10.72 (1.86)	10.01 (1.73)	<0.001
HCT (hematocrit) *TWA (sd)	31.82 (5.32)	30.04 (5.10)	<0.001
RDW (red blood cell distribution width)*TWA (sd)	15.11 (2.07)	16.30 (2.46)	<0.001
RBC (red blood cells)*TWA (sd)	3.57 (0.64)	3.34 (0.61)	<0.001
MCHC (mean corpuscular hemoglobin concentration) *TWA (sd)	33.69 (1.56)	33.26 (1.71)	<0.001
MCH (mean corpuscular hemoglobin) *TWA (sd)	30.15 (2.52)	30.20 (2.78)	0.708
Lymphocytes *TWA (sd)	12.33 (11.67)	11.91 (13.83)	0.671
WBC (white blood cells)*TWA (sd)	11.51 (7.89)	13.04 (13.10)	<0.001
Eosinophiles *TWA (sd)	1.01 (1.96)	1.12 (2.95)	0.493
Basophiles *TWA (sd)	0.25 (0.36)	0.20 (0.30)	0.149
Bands* TWA (sd)	5.70 (8.72)	7.54 (9.62)	0.073
Monocytes* TWA (sd)	4.55 (3.87)	4.39 (3.50)	0.622
PLT (platelets concentration)*TWA (sd)	221.39 (114.47)	178.93 (113.13)	<0.001
MCV (mean corpuscular volume) *TWA (sd)	89.57 (6.61)	90.86 (7.35)	<0.001
BE (base excess) *TWA (sd)	-0.42 (4.52)	-3.54 (5.04)	<0.001
Neutrophils *TWA (sd)	79.06 (15.55)	76.11 (19.06)	0.025

Abbreviations Used: *sd: standard deviation, *TWA: Time-weighted average, Na: sodium, K: potassium, Cl: chloride, Bicarb: sodium bicarbonate, firstBicarb: first bicarb recorded within the first 24 hours, minBicarb: minimal from all bicarb values recorded within the first 24 hours, SCr: serum creatinine, PTH: parathyroid hormone, Phos: phosphorus, BUN: blood urea nitrogen, SBP: systolic blood pressure, HR: heart rate.*

Table 1: Characteristic of the study population, stratified by the presence or iRRT outcome.

Calculating SOFA, OASIS and APS III scores

Sequential Organ Failure Assessment (SOFA) [15], Oxford Acute Severity of Illness Score (OASIS) [16] and Acute Physiology Score III (APS III) [17] scores were calculated using source code from the official MIMIC 3 code repository [18].

Definition of incident renal replacement therapy (iRRT)

The outcome, iRRT within eight days of ICU stay, was defined as any dialysis-related event charted in EMR within the seven days following the first 24 hours of ICU admission. Supplementary Table 3 lists the events used to annotate iRRT.

Scanning database for charted iRRT related events

In order to find RRT events, the database, and specifically its table "D_ITEMS" was scanned for any data point, with *category*, *label* or *abbreviation* descriptive field including any of the following phrases: "dialy","renal rep","crrt","cvv","ultrafilt" and "uf". This search returned 72 items. We manually selected those related to RRT. Supplementary Table 3 lists all annotating RRT events. Although very unlikely, we understand that it is possible, that using above described approach, not all RRT events were found. Renal replacement therapy, on the other hand, is a complicated procedure, usually requiring multiple various events being charted in EMR. We assumed, therefore, that at least one of the events listed in Supplementary Table 3 would be found during the database scanning. With this approach, out of all 18379 ICU stays, we found 341 (1.85%) ICU admissions with iRRT started between 24 hours and eight days-time points after ICU admission.

Artificial intelligence (AI) algorithm

A significant amount of data collected during single ICU stay make the use of AI especially appealing. Random Forests (RF) [19] is one of the most popular and accurate methods of contemporary AI science, yet simple to use. Random Forest creates a set of independent decision trees. Each decision tree is "grown" based on decision points (i.e., gender: male or female) efficiently dividing a dataset into outcome groups. In our case, a group of patients who received iRRT and a group of patients who did not receive iRRT. Once all "trees are grown," they are subsequently applied to predict the outcomes for patients in testing

dataset. The final prediction is calculated as an average from predictions returned by all trees.

Searching for optimal machine learning model and optimal parameters to train it

We programmed grid search to find optimal parameters to train AI model, with selected combinations of values for *ntrees* (from 50, 100, 120, 200, 400, 500), number of cross-validation runs (from 10, 80, 130, 180, 200, 400), *max_depth* (from 3,4,5,6) and flag *balance_classes* set to TRUE or FALSE. The maximal mean area under the curve (AUC) of receiver-operator characteristics curve (ROC) for all cross-validation runs, was seen with this set of parameters: *ntrees* set to 100, *max_depth* set to 5 and *balance_classes* set to FALSE, and these parameters were used for further experiments.

Data analysis

The main experiment was programmed, with parameters found as explained in section 2.8, to perform 400 cross-validation runs. For each of runs, 92% of 18359 cases were randomly chosen to constitute a training data set, and AI was trained to predict the outcome of iRRT. Using trained AI RF model, predictions were made for the remaining 8% of the cases, constituting the testing dataset. These predictions, and predictions made by SOFA, OASIS and APS III scores, were validated against the known outcome of iRRT. The area under the curve (AUC) of receiver operating characteristics (ROC) curve was calculated and recorded for each of cross-validation runs. Figure 1 illustrates the process. Additionally, using an internal functionality of RF to measure the intrinsic importance of variables, the importance of each variable, averaged across all 400 iterations, was saved.

Result

Overall the incidence of RRT in 400 testing datasets was 27.2 [95% CI 26.7-27.7] cases. For 400 AI-models, mean AUC of ROC was 0.885 [95% CI 0.882-0.889]. Mean AUC for SOFA, APSIII and OASIS scores predicting iRRT respectively was 0.821 [95% CI 0.818-0.825], 0.809 [95% CI 0.806-0.813], and 0.697 [95% CI 0.693-0.702]. Figure 2 illustrates results of exemplary run.

Comparing AI-model vs. each of SOFA, APSIII and OASIS with Wilcoxon signed-rank paired revealed statistically significant difference ($p<0.001$). Table 2 lists the 20 most essential predictors, as returned by AI-model. In a similar experiment, with the understanding that 679 variable model, would not be feasible for manual data entry, we built additional "mini-AI-model" using only 20 most important variables found in the main experiment. This "mini-AI-model's" mean AUC of ROC was 0.86 [95% CI 0.856-0.864] and remained superior to SOFA, APSIII and OASIS severity scores as tested with Wilcoxon signed-rank paired test ($p<0.001$).

Discussion

Our study proves the concept, that data automatically and routinely collected in electronic medical records (EMR) and typically not analyzed to predict renal outcomes, can be used to predict the incident renal replacement therapy in ICU population. The legacy of our study design and the definition of iRRT bring additional advantage; our model is predicting actual incident renal replacement therapy performance, not only indication for RRT. It predicts patients who will survive until iRRT is administered vs. patients who maybe had iRRT indicated, but after all, had not undergone iRRT due to their pre-emptive death, their transfer to different hospital for continued care, lack of consent for iRRT or contraindication to receive iRRT. We postulate that our approach can be instituted in hospitals' EMRs, to analyze the EMR's data in background and real-time, to predict the actual likelihood of administration of RRT in close future. These predictions can be further used to alarm healthcare team, possibly lead to nephrology consult, with preemptive planning for RRT access options. Nursing staffing could be efficiently planned, initiation of prophylactic measures like nephrotoxins or high potassium IV fluids and diet avoidance could be instituted, or if necessary, patients could be transferred to the hospital, where RRT is available. Models like ours can function as add-on alerts to EMR. Additionally, trained locally, on the hospital's database and

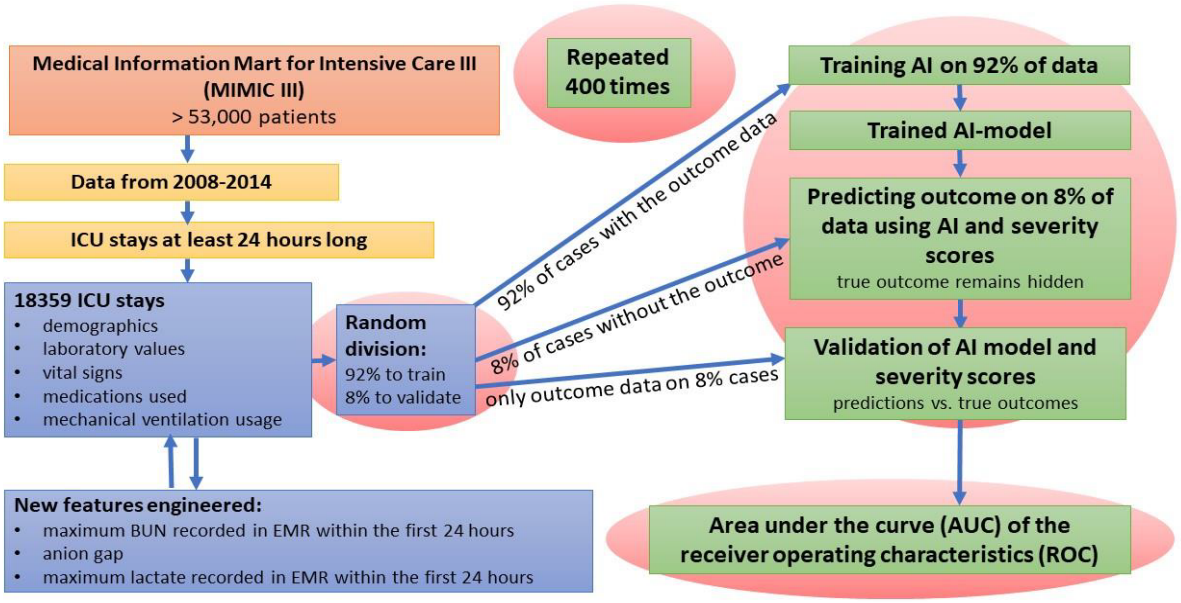

Figure 1: Scheme of the analysis. Encircled actions where repeated 400 times.

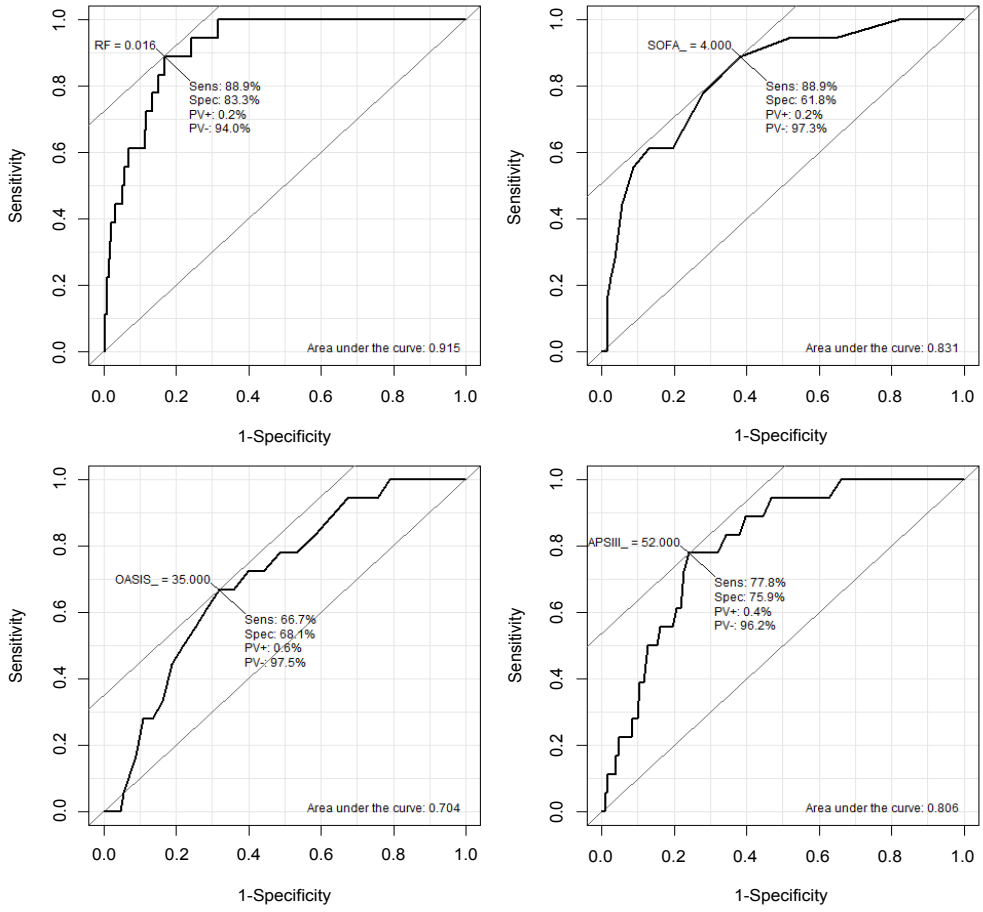

Figure 2: Receiver-operator characteristics curves (ROC) for one exemplary run are shown for random forest (RF) model (top-left), SOFA (top-right), OASIS (bottom-left) and APSIII (bottom-right). *Abbreviations used: Sens for sensitivity, Spec for specificity, PV+ for positive predictive value and PV- for negative predictive value.*

Variable name	Explanation	Average% importance
LastSCr	Last recorded SCr within 24 hours of ICU stay.	0.033
SCr/UOpt/minSats/minpH	TWA SCr. Minimal values for Sats and pH taken.	0.027
MaxSCr	Maximal recorded SCr within 24 hours of ICU stay.	0.026
SCr time-weighted average	TWA SCr.	0.025
SCr/UOpt	TWA serum creatinine.	0.024
meanSCr	The simple arithmetic mean of all recorded SCr within first 24 hours.	0.023
minSCr	Minimal recorded SCr within 24 hours of ICU stay.	0.021
deltaSCr/UOpt	(Last seen SCr – first seen SCr) divided by urine output.	0.02
deltaSCr/UOpt/minSats/minpH	As above, then divided by both minimal values for Sats and pH.	0.02
delta SCr	Last seen SCr – first seen SCr.	0.02
delta SCr or ZERO	delta SCr or zero, if only one SCr available within the 24 first hours.	0.019
Phos/UOpt	TWA Phos / urine output. Usually high in chronic kidney disease (CKD).	0.017
LastSCr	First recorded SCr within 24 hours of ICU stay.	0.016
LastPhos	Last recorded Phos within 24 hours of ICU stay.	0.014
TWA Phos	Time-weighted average phosphorus. May be high in CKD.	0.013
TWA K/UOpt	May be high in acute kidney injury (AKI).	0.013
UOpt	May be low both in AKI and CKD.	0.013
meanPhos	Simple arithmetic means of all recorded Phos values within the first 24 hours.	0.012
UOpt/TWA weight	May be low both in AKI and CKD.	0.011
maxPhos	Maximal of all recorded Phos values within the first 24 hours.	0.011
Abbreviations Used: *sd: standard deviation, TWA: Time-weighted average, K: potassium, Bicarb: sodium bicarbonate, SCr: serum creatinine, PTH: parathyroid hormone, Phos: phosphorus, UOpt: urine output recorded within first 24 hours of ICU stay, minpH: minimal value of pH recorded within first 24 hours, minSats: minimal value of O2 saturation recorded within first 24 hours*		

Table 2: Top-20 most important predictors, as found by AI-model and used for mini-AI-model.

past patients, are tuned into the hospital environment, resources, population, and physicians' styles of practice, which offers the best possible prediction for that given scenario.

Striking is the accuracy of our model to predict iRRT beyond 24 hours, only on the first 24 hours of data. In clinical reality, many events, which lead to iRRT can happen in ICU, after the first 24 hours. These events include contrast administration, volume overload, need for surgery, hypotension, or worsening hyperkalemia. We showed that the first 24 hours of patient-related data and events, determine if the patient will receive the iRRT. That reinforces the importance of the early introduction of prophylaxis and treatment of acute kidney injury and its consequences. To our knowledge, this is the first ever attempt to use machine learning or AI and such engineered features to predict iRRT.

Another advantage of our model, inherent to h2o.ai libraries design, is that it handles missing data and infinite values without interaction from the operator. No assumption is necessary to handle missing data. AI model leverages their value, as can treat missing data, as they were "missing for a reason".

AI-model built on 679 is not feasible for use in clinical medicine without the support of computers and automatic EMR data retrieval. As a solution to this, we propose the "mini-AI-model" based only on 20 predictors. Additional limitations are inherent to the retrospective calculation of SOFA, OASIS and APS III scores.

Conclusion

Our research was the first know to authors, estimating the usefulness of SOFA, APSIII and OASIS severity scores to predict the need for iRRT. Surprisingly, even though these scores can be calculated manually with relatively low effort, their performance is not to be ignored. To our knowledge, this is the first ever evaluation of the applicability of these severity scores to predict iRRT.

Trends and specific values of serum creatinine, pH, phosphorus, oxygen saturation, urine output, patient's weight and potassium were found as the most important by AI-model. These resemble our physicians' approach to the problem of initiation of iRRT in ICU: we look at the acid-base balance, potassium level, volume status and urine output.

More research, best with external validation of our approach, is warranted to assess clinical applicability of our approach.

We are eager to share the portions of the source code for all analyses, to the degree allowed by MIMIC database administrators to promote and encourage other investigators to use MIMIC 3 and overall use of machine learning methods in medicine and collaborations. The data has to be requested form MIMIC 3 administrators.

Author Contribution

LK designed the study, extracted and analyzed the data, wrote the manuscript. S.A. co-designed the study and critically revised manuscript. All authors approved the final version of the manuscript.

Conflict of Interest

Lukasz Kiljanek MD has no conflict of interest to declare, Sandeep Aggarwal MD has no conflict of interest to declare.

References

1. Case J, Khan S, Khalid R, Khan A (2013) Epidemiology of Acute Kidney Injury in the Intensive Care Unit. Crit Care Res Pract 479730.
2. Nisula S, Kaukonen KM, Vaara ST, Korhonen AM, Poukkanen M, et al. (2013) Incidence, risk factors and 90-day mortality of patients with acute kidney injury in Finnish intensive care units: the FINNAKI study. Intensive Care Med 39: 420-428.
3. Allyn J, Allou N, Augustin P, Philip I, Martinet O, et al. (2017) A Comparison of a machine learning model with EuroSCORE II in predicting mortality after elective cardiac surgery: A decision curve analysis. PLoS One 12: e0169772.
4. Venkatesh A, Jacob SG (2016) Prediction of credit-card defaulters: A comparative study on performance of classifiers. Int J Comp Appl 145: 7.
5. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, et al. (2016) MIMIC-III, a freely accessible critical care database. Sci Data 24: 160035.
6. Henry KE, Hager DN, Pronovost PJ, Saria S (2015) A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med 7: 299ra122.
7. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, et al. (2015) Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): A population-based study. Lancet Respir Med 3: 42-52.
8. The pgAdmin Development Team, pgAdmin, 1.4 ed, 2013-2017.
9. R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
10. Revolution Analytics and Stephen Weston. doSNOW: Foreach Parallel Adaptor for the 'snow' Package. R package version 1.0.14.
11. Csárdi G, FitzJohn R (2016) Progress: Terminal Progress Bars, R package version 1.1.2 ed.
12. Yoshida K, Bohn J (2015) tableone: Create "Table 1" to Describe Baseline Characteristics. R package version 0.7.3.
13. Dowle M, Srinivasan A (2017) data.table: Extension of `data.frame`. R package version 1.10.4 ed.
14. Yoshida K (2014) Introduction to tableone.
15. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, et al. (1996) The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive Care Med 22: 707-710.
16. Alistair JEW, Kramer AA, Clifford GD (2013) A new severity of illness scale using a subset of acute physiology and chronic health evaluation data elements shows comparable predictive accuracy. Critical Care Med 41: 1711-1718.
17. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, et al. (1991) The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 100: 1619-1936.
18. MIMIC 3 database administrators (2017) MIMIC 3 database code repository.
19. Breiman L (2001) Random Forests. Machine Learning 45: 5-32.