# Real-Time Prediction of AKI Among Middle-Aged and Older in ICU: A Retrospective and Machine Learning Study

**Hui Wang**
School of Automation Central South University
**Fuxing Deng**
Xiangya Hospital Central South University
**Buyao Zhang**
Xiangya Hospital Central South University
**Shuangping Zhao** ( zhshping@csu.edu.cn )
Xiangya Hospital Central South University https://orcid.org/0000-0003-1258-8852

**Research**

**Keywords:** AKI, mortality, prediction, neural network

**Posted Date:** August 17th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-57905/v1

# Abstract

### Background

Acute Kidney Injury (AKI), a major public health problem is responsible for two-thirds of intensive care unit patients' cost, and aging is an independent risk factor for AKI and its associated mortality and morbidity. The early recognition of AKI helps ICU caregivers to guide fluid treatment and titrate the dosing of the nephrotoxic drug. Therefore, it is desirable to build models to predict their position. The study is to build models based machine learning to predict AKI stage after 24 hours and 48 hours among middle-aged and older patients respectively in ICU.

### Methods and Findings

We used two real-world databases to build and test models. The Medical Information Mart for Intensive Care (MIMIC-III v1.4) database for training, funded by National Institutes of Health (NIIH) built by the Computational Physiology Laboratory of MIT, Beth Israel Dikon Medical Center, and Philips Medical. The eICU Collaborative Research Database (eICU-CRD v 2.0) for the test is open-access, de-identified data sets of patients admitted to ICUs. 26316 patients in the overall cohort were generally older (median age ranging from 57 to 79) and 54% were male. Here we present three models, using the support vector machine (SVM), Long short-term memory (LSTM), and convolutional LSTM ConvLSTM respectively. the ConvLSTM model had the best performance in the test data set, and it has good ability and surpasses any previous model to predict whether older patients have AKI or not. The area under the receiver operating characteristic curve (AUC) of 24-hour prediction AKI is 99.79%, 48-hours AKI 99.43% during the hospital. we demonstrate that deep learning can handle lots of variables which may be predictors and that the algorithm achieved robust and excellent performance.

### Conclusions

To our knowledge, this study is the first to use large-scale data collected from electronic health record EHR to prove the contribution of big data and deep learning methods to the real-time prediction of AKI prognosis in middle-aged and elderly patients. The model performance is better than any previous models. This work provides novel evidence to change clinical practice and precise personalized interventions.


# Introduction

Acute Kidney Injury (AKI), a major public health problem, is defined by rapidly ascending value in serum creatinine or a descending urine output according to Kidney Disease: Improving Global Outcomes (KDIGO) guidelines. It is responsible for two-thirds of intensive care unit patients' cost and is associated with increased length of intensive care unit (ICU) stay as well as high mortality[1, 2]. The early recognition of AKI helps ICU caregivers to guide fluid resuscitation and titrate the dosing of the nephrotoxic drug. Recently, population aging has emerged as a global phenomenon in the wake of the now virtually universal decline in fertility and increases in life expectancy. In China, data from the national bureau of statistics demonstrate that the proportion of elderly people increases between 2009 to 2018. Some investigators found that aging is an independent risk factor for AKI and its associated mortality and morbidity[3–6]. Since AKI patients among middle-aged and older in ICU have high mortality and progress easily into Chronic kidney disease (CKD) after discharging ICU[7–13], so it is desirable to build models to predict their position.

In recent years, with the rapid development of deep learning technology, artificial intelligence (AI) and machine learning (ML) have been used to handle with a variety of problems[14], ranging from computer vision, gaming, high-energy physics, drug-design[15] and bioinformatics[16]. Besides, many deep learning methods are also applied to the field of medical image[17] analysis, but to building deep a neural network model by using lab items and vitals is relatively limited. Recently, some investigations used several AI/ML methods to build models, which could predict AKI 61.8(32.5) hours faster than the Kidney Disease and Improving Global Disease Outcomes (KDIGO) criteria for burn and non-burned trauma patients using NGAL, creatinine, NT-proBNP, and UOP, and plasma creatinine[18]. Another study using machine learning techniques is to predict AKI risk of patients undergoing percutaneous coronary intervention (PCI)[19]. We also have found the proposed nomogram effectively predicted AKI risk in sepsis patients admitted to the intensive care unit in the first 24 h[20], but there is a lack of predicting AKI models for middle-aged and older patients in ICU.

To explore the deep learning in predicting the AKI stage with the increase of hospitalization time among middle-aged and older patients in ICU, we sought to build deep learning models based on recurrent neural network (RNN). Deep learning methods may provide advantages for AKI stage identification, because they can be trained to predict AKI for in advance of onset, can maintain concurrently high sensitivity and specificity, and can be customized to specific populations for increased accuracy. In this study, we present the development and validation of deep neural network models based on basic and clinical information of patients. One model is to predict AKI stage after 24 hours and 48 hours, which can be refreshed in dynamic real-time among middle-aged and older patients in ICU.

## Method

We used patients' personal basic information and daily clinical examination data from MIMIC and eICU to build three prediction models using deep neural network, through data preprocessing and the establishment of feature engineering. This study is reported in accordance with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement [https://www.equator-network.org/reporting-guidelines/tripod-statement/].

# Data Source and Study Population

This retrospective, data mining study included more than 40000 ICU patients hospitalized between 2001 to 2012. Deidentified patients' information was abstracted from the Medical Information Mart for Intensive Care (MIMIC-III v1.4) database, funded by National Institutes of Health (NIIH)[21], which built by the Computational Physiology Laboratory of MIT, Beth Israel Dikon Medical Center, and Philips Medical. The eICU Collaborative Research Database (eICU-CRD v 2.0)[22], which covers 200859 ICU admission between 2014 and 2015 of 139367 patients at 208 U.S. hospitals, is also open-access, de-identified data sets of patients admitted to ICUs. Both databases are maintained by the Laboratory for Computational Physiology at the Massachusetts Institute of Technology. They include hourly vital recordings from bedside monitors, diagnoses via International Classification of Diseases, Ninth Revision (ICD-9) codes, records of basic personal characteristic information, and other clinical data, collected during routine medical care. The information data of relevant patients from the MIMIC and eICU database through PostgreSQL software (version 9.6).

The cohort included patients ≥ 45 years old and diagnosed stage of Acute Kidney Injury according to the KDIGO criteria. Figure 1 shows the patient selection and reasons for exclusion. In eICU database, we excluded 3550

secondary (or greater) admissions for patients to avoid repeated measures, 3002 patients whose length of stay in ICUs are less than 24 hours since no label of those patients can be to predict, 2241 admissions with missing data and developing AKI in ICU first day. In MIMIC-III data, we also excluded 6968 secondary (or greater) admissions for patients to avoid repeated measures, 6835 patients whose length of stay in ICUs are less than 24 hours since no label of those patients can be to predict, 22340 admissions with missing data and developing AKI in ICU first day. The final cohort contained 26316 patients to predict AKI (6130 patients in eICU as test sets, 20186 patients in MIMIC-III as training sets).

# Study outcome

The outcome for the diagnosis model is the stage of AKI, defined by the Kidney Disease: Improving Global Outcomes (KDIGO) criteria: stage 1 means increasing in serum creatinine ≥ 26.4 μmol//L or increasing to ≥ 1.5-fold to twofold from baseline, stage 2 means > twofold to threefold from baseline and stage 3 means > threefold from baseline or serum creatinine ≥ 354 μmol/L with an acute increase of at least 44 μmol/L. Individuals who receive RRT are considered to have met the criteria of AKI III regardless of their serum creatinine value. Stage 0 means AKI is not predicted here. Serum creatinine baseline according to KDIGO standards for all patients.

# Data preprocessing

Predictors included laboratory examinations, vital signs, the number of using vasopressors' times, urine output, whether to use mechanical ventilation, and whether to use sedatives. Patients' basic information, such as age, height, weight, and race, was incorporated into prediction models. All clinical data from the first 7 days of ICU admission for each patient were extracted for building models. Variables that were labeled by the 'DataExplore' package in R were removed. Then we used the 'mice' package to impute missing values by iterated 50 times, and the method of 'mice' selected 'pmm' (predictive mean matching). All numerical items were normalized to the [0,1] range, and some categorical variables have been encoded by OneHot way. Then query the sample for screening, which screening method is as follows: If one patient's initial AKI stage is not 0 at every 24 hours after admission, it means the patient data is discarded. If one patient's initial AKI stage is 0, the stage of AKI of the next day (after 24 hours) or the next two days (after 48 hours) are used as the prediction labels, and the labels are saved with the data for training. All the code that extracts the data and processes the missing values is in the Github (https://github.com/ZEROICEWANG/Older-aki).

# Algorithm development and validation

We used three machine learning algorithms to train the model, support vector machine, long short-term memory (LSTM), and convolution long short-term memory (ConvLSTM) algorithm. Support vector machine (SVM) is a classical supervised machine learning model, which can realize nonlinear classification by the kernel function. In this experiment, the Radial Basis Function (RBF) kernel is used. Long Short-Term Memory (LSTM) network is a kind of time cycle neural network, which is specially designed to solve the long-term dependence problem of general RNN. The ConvLSTM model for predicting the AKI stage in the near future contains 3 parts, 1) Convolution pooling layers, this layer is composed of two layers of convolution, and two layers of pooling. The first convolution layer has 64 convolution kernels with a size of 3 * 3 and the stride step is 1. The first pooling layer has a size of 2 * 2 and a step size of 2. The second convolution layer has 128 convolution kernels with a size of 5 * 5 and a step size of 1. The second pooling layer has a size of 3 * 3 and a step size of 3. In addition to the convolution and pooling layers, adding a Batch-Normalization layer is to effectively improve the convergence speed in addition to avoiding overfitting. 2) Convolutional long short-term memory (ConvLSTM) layers, after convolution pooling the

middle layer, the data is reshaped and connected to ConvLSTM. This part consists of two layers of ConvLSTM, each layer has 256 convolution kernels, the size of the convolution kernel is 3 * 3, and the step size is 1. 3) Fully connected layers, after ConvLSTM, flatten the data with one flatten layers, and finally connect with the fully connected layer, The size of the first fully connected layer is 2560, the second fully connected layer is 2048, and the third fully connected layer is 4. Network parameters U, V, and W were initialized randomly and were successively refined as the model learned to recognize the various clinical complications from historic episodes, the so-called training data. The optimization selected was Adam,and batch size was 500, and the number of epochs was 100. The schematic diagram of the model is shown in Fig. 2. All RNN models were developed and applied using Tensorflow 1.7. All the modeling code can be found in the Github (https://github.com/ZEROICEWANG/Older-aki).

# Statistical Analysis

To evaluate the performance of those models, several different methods were conducted. The prediction quality was evaluated by precision, recall and F1-score metrics defined as,

$$\text{Precision} = \frac{TP}{TP+FP},$$

$$\text{Recall} = \frac{TP}{TP+FN},$$

$$\text{F1 score} = 2 * \frac{precision*recall}{precision+recall}$$

Where TP represents true positive (i.e., patients with the grade 1 AKI is predicted correctly as that AKI category), FP false positives (patients with no-AKI is predicted as from the category of AKI), and FN false negatives (patients with AKI from different grade is predicted as no AKI). We also used the area under the receiver operating characteristic curve (AUC, also known as the c-statistic) to measure model discrimination.

Between the two groups was analyzed statistically and was using the two-sample t-test, Pearson X2 test, or the Mann-Whitney-Wilcoxon U test, which these values were calculated in R by ‘CBCgrps’ package. ALL data were analyzed using R the statistical package (version 3.6.1). The statistical significance level for all the tests was set at a P-value < 0.05.

# Results

A description of the study population is displayed in Table 1. In total, 20186 individuals from MIMIC are used to build models. The mean age (IQR) 67.8 (57.6, 78.4) in MIMIC sets and 69.0 (59.0, 79.0) in eICU sets are demonstrated that this study focuses on middle-aged and older patients. Although some P values are statistically significant, the interquartile regions of all variables overlap.

Table 1
Baseline demographic and laboratory characteristics of patients at first admission day between patients of eICU and MIMIC III.

| Variables | Total (n = 26316) | eICU (n = 6130) | MIMIC (n = 20186) | P value |
|---|---|---|---|---|
| Age, Median (IQR) | 68 (58, 78.38) | 68 (59, 79) | 67.75 (57.64, 78.36) | < 0.001 |
| Gender, n (%) | | | | < 0.001 |
| Female | 11610 (44) | 2984 (49) | 8626 (43) | |
| Male | 14706 (56) | 3146 (51) | 11560 (57) | |
| Height, Median (IQR) | 170 (160.02, 177.8) | 168 (161.3, 177.8) | 170 (160.02, 177.8) | 0.344 |
| Weight, Median (IQR) | 75 (63.4, 89) | 78.5 (65, 95.3) | 74.5 (63, 87.2) | < 0.001 |
| Anion Gap, Median (IQR) | 14 (11, 16) | 11 (8, 14) | 14 (12, 17) | < 0.001 |
| Bicarbonate, Median (IQR) | 25 (23, 28) | 25 (22, 28) | 25 (23, 28) | < 0.001 |
| Creatinine, Median (IQR) | 1 (0.72, 1.35) | 1.04 (0.77, 1.58) | 1 (0.7, 1.3) | < 0.001 |
| Chloride, Median (IQR) | 106 (102, 110) | 104 (100, 109) | 107 (103, 111) | < 0.001 |
| Glucose, Median (IQR) | 148 (118, 190) | 133 (107, 175) | 153 (123, 194) | < 0.001 |
| Hematocrit, Median (IQR) | 34.7 (30.67, 39) | 32.9 (28.4, 37.5) | 35 (31.3, 39) | < 0.001 |
| Hemoglobin, Median (IQR) | 11.6 (10.1, 13) | 10.8 (9.2, 12.4) | 11.8 (10.4, 13.2) | < 0.001 |
| Platelet, Median (IQR) | 217 (162, 288) | 192 (143, 251) | 226 (170, 298) | < 0.001 |
| Potassium, Median (IQR) | 4.4 (4, 4.9) | 4.1 (3.8, 4.6) | 4.4 (4, 5) | < 0.001 |
| Sodium, Median (IQR) | 140 (137, 142) | 139 (136, 141) | 140 (138, 142) | < 0.001 |
| BUN, Median (IQR) | 20 (14, 32) | 23 (15, 38) | 20 (14, 31) | < 0.001 |
| WBC, Median (IQR) | 11.4 (8.3, 15.41) | 10.41 (7.4, 14.61) | 11.6 (8.5, 15.6) | < 0.001 |

Abbreviations:

BUN, blood urea nitrogen; WBC, white blood cell count; SPO2, Oxygen saturation; LOS in ICU, length of stay in ICU (day).

| Variables | Total (n = 26316) | eICU (n = 6130) | MIMIC (n = 20186) | P value |
|---|---|---|---|---|
| Heart rate, Median (IQR) | 100 (88, 114) | 101 (88, 116) | 100 (88, 114) | < 0.001 |
| Systolic pressure, Median (IQR) | 148 (133, 165) | 148.5 (131, 168) | 148 (134, 164) | 0.41 |
| Diastolic blood pressure, Median (IQR) | 82 (72, 94) | 85 (73, 99) | 81 (72, 93) | < 0.001 |
| Mean arterial pressure, Median (IQR) | 101 (91, 114) | 102 (90, 116) | 101 (91, 113) | 0.009 |
| Respiratory rate, Median (IQR) | 26 (23, 31) | 27 (23, 34) | 26 (23, 30) | < 0.001 |
| Temperature Median (IQR) | 37.3 (36.94, 37.83) | 37.2 (36.9, 37.6) | 37.39 (37, 37.9) | < 0.001 |
| SPO2, Median (IQR) | 100 (99, 100) | 100 (99, 100) | 100 (100, 100) | < 0.001 |
| LOS in ICU, Median (IQR) | 2.27 (1.54, 4.06) | 2.29 (1.62, 3.96) | 2.27 (1.52, 4.09) | 0.438 |
| Congestive Heart Failure, n (%) | | | | < 0.001 |
| no | 22405 (86) | 5364 (91) | 17041 (84) | |
| yes | 3683 (14) | 538 (9) | 3145 (16) | |
| Cardiac Arrhythmias, n (%) | | | | < 0.001 |
| no | 21593 (83) | 5117 (87) | 16476 (82) | |
| yes | 4495 (17) | 785 (13) | 3710 (18) | |
| Valvular Disease, n (%) | | | | < 0.001 |
| no | 24839 (95) | 5877 (100) | 18962 (94) | |
| yes | 1249 (5) | 25 (0) | 1224 (6) | |
| Pulmonary Circulation, n (%) | | | | < 0.001 |
| no | 25196 (97) | 5787 (98) | 19409 (96) | |
| yes | 892 (3) | 115 (2) | 777 (4) | |
| Peripheral Vascular, n (%) | | | | < 0.001 |
| no | 24132 (93) | 5842 (99) | 18290 (91) | |

Abbreviations:

BUN, blood urea nitrogen; WBC, white blood cell count; SPO2, Oxygen saturation; LOS in ICU, length of stay in ICU (day).

| Variables | Total (n = 26316) | eICU (n = 6130) | MIMIC (n = 20186) | P value |
|---|---|---|---|---|
| yes | 1956 (7) | 60 (1) | 1896 (9) | |
| Hypertension, n (%) | | | | < 0.001 |
| no | 23667 (91) | 5429 (92) | 18238 (90) | |
| yes | 2421 (9) | 473 (8) | 1948 (10) | |
| Paralysis, n (%) | | | | < 0.001 |
| no | 25436 (98) | 5890 (100) | 19546 (97) | |
| yes | 652 (2) | 12 (0) | 640 (3) | |
| Chronic pulmonary, n (%) | | | | < 0.001 |
| no | 21372 (82) | 5277 (89) | 16095 (80) | |
| yes | 4716 (18) | 625 (11) | 4091 (20) | |
| Diabetes Uncomplicated, n (%) | | | | < 0.001 |
| no | 21716 (83) | 5724 (97) | 15992 (79) | |
| yes | 4372 (17) | 178 (3) | 4194 (21) | |
| Diabetes Complicated, n (%) | | | | < 0.001 |
| no | 25062 (96) | 5902 (100) | 19160 (95) | |
| yes | 1026 (4) | 0 (0) | 1026 (5) | |
| Hypothyroidism, n (%) | | | | < 0.001 |
| no | 23906 (92) | 5836 (99) | 18070 (90) | |
| yes | 2182 (8) | 66 (1) | 2116 (10) | |
| Liver Disease, n (%) | | | | < 0.001 |
| no | 24794 (95) | 5822 (99) | 18972 (94) | |
| yes | 1294 (5) | 80 (1) | 1214 (6) | |
| Lymphoma, n (%) | | | | < 0.001 |

Abbreviations:

BUN, blood urea nitrogen; WBC, white blood cell count; SPO2, Oxygen saturation; LOS in ICU, length of stay in ICU (day).

| Variables | Total (n = 26316) | eICU (n = 6130) | MIMIC (n = 20186) | P value |
|---|---|---|---|---|
| no | 25765 (99) | 5881 (100) | 19884 (99) | |
| yes | 323 (1) | 21 (0) | 302 (1) | |
| Metastatic Cancer, n (%) | | | | < 0.001 |
| no | 25118 (96) | 5849 (99) | 19269 (95) | |
| yes | 970 (4) | 53 (1) | 917 (5) | |
| Solid Tumor, n (%) | | | | 0.036 |
| no | 25246 (97) | 5686 (96) | 19560 (97) | |
| yes | 842 (3) | 216 (4) | 626 (3) | |
| Rheumatoid Arthritis, n (%) | | | | < 0.001 |
| no | 25428 (97) | 5890 (100) | 19538 (97) | |
| yes | 660 (3) | 12 (0) | 648 (3) | |
| Coagulopathy, n (%) | | | | < 0.001 |
| no | 23788 (91) | 5711 (97) | 18077 (90) | |
| yes | 2300 (9) | 191 (3) | 2109 (10) | |
| Weight Loss n (%) | | | | < 0.001 |
| no | 25161 (96) | 5881 (100) | 19280 (96) | |
| yes | 927 (4) | 21 (0) | 906 (4) | |
| Fluid Electrolyte n (%) | | | | < 0.001 |
| no | 19925 (76) | 5358 (91) | 14567 (72) | |
| yes | 6163 (24) | 544 (9) | 5619 (28) | |
| Blood Loss Anemia, n (%) | | | | < 0.001 |
| no | 25610 (98) | 5898 (100) | 19712 (98) | |
| yes | 478 (2) | 4 (0) | 474 (2) | |
| Deficiency Anemias, n (%) | | | | < 0.001 |

Abbreviations:

BUN, blood urea nitrogen; WBC, white blood cell count; SPO2, Oxygen saturation; LOS in ICU, length of stay in ICU (day).

| Variables | Total (n = 26316) | eICU (n = 6130) | MIMIC (n = 20186) | P value |
|---|---|---|---|---|
| no | 22294 (85) | 5872 (99) | 16422 (81) | |
| yes | 3794 (15) | 30 (1) | 3764 (19) | |
| Alcohol Abuse, n (%) | | | | < 0.001 |
| no | 24728 (95) | 5780 (98) | 18948 (94) | |
| yes | 1360 (5) | 122 (2) | 1238 (6) | |
| Drug Abuse, n (%) | | | | < 0.001 |
| no | 25617 (98) | 5899 (100) | 19718 (98) | |
| yes | 471 (2) | 3 (0) | 468 (2) | |
| Psychoses, n (%) | | | | < 0.001 |
| no | 25303 (97) | 5878 (100) | 19425 (96) | |
| yes | 785 (3) | 24 (0) | 761 (4) | |
| Depression, n (%) | | | | < 0.001 |
| no | 24270 (93) | 5868 (99) | 18402 (91) | |
| yes | 1818 (7) | 34 (1) | 1784 (9) | |
| Abbreviations: | | | | |
| BUN, blood urea nitrogen; WBC, white blood cell count; SPO2, Oxygen saturation; LOS in ICU, length of stay in ICU (day). | | | | |

A total of 26316 participants were included. Figure 1 shows the patient selection and reasons for exclusion. The distribution of missing data values is in the supplementary 'missing_eICU' and 'missing_MIMIC' file. The average age was 68 years with 44% women in all data, and the average height was 170.0 (160.0, 177.8) and 168 (161.3, 177.8), average weight 74.5 (63.0, 87.2) and 78.5 (65.0, 95.3) in training sets and test sets respectively. The comparison results of some basic laboratory inspection items and basic vitals indexes are also shown in Table 1. Although some P values are statistically significant, the interquartile regions of all variables overlap.

# Model performance

There are some indexes for the evaluation of those models using the test data set (n = 6130). Compared to SVM model (AUC of 24-hour prediction: 0.928, AUC of 48-hour prediction: 0.937) and LSTM model (AUC of 24-hour prediction: 0.956, AUC of 48-hour prediction: 0.948), we observed that ConvLSTM model had the largest AUC (AUC of 24-hour prediction: 0.998, AUC of 48-hour prediction: 0.994) (supplementary AUC). Accuracy, Precision, Recall, F1, PPV, NPV, Sensitivity, and Specificity are the prediction results for the test set (Table 2). Figure 3 illustrates the calibration curve for the ConvLSTM model. The sensitivity and specificity of the 24-hour AKI classification based ConvLSTM model were 95.22%, 98.00%, positive predictive value was 94.64%, and the negative predictive value

was 98.13%. In the 48-hour AKI prediction model, the sensitivity, specificity, and positive predictive value were 90.90%, 96.83%, 91.50%, and 97.19% respectively.

Table 2
Performance of the three models

| 24 hour prediction Model | Accuracy | Precision | Recall | F1 | PPV | NPV | Sensitivity | Specificity |
|---|---|---|---|---|---|---|---|---|
| SVM | 0.82457 | 0.93883 | 0.68926 | 0.77098 | 0.93883 | 0.94941 | 0.68926 | 0.89334 |
| LSTM | 0.77234 | 0.73457 | 0.68264 | 0.70412 | 0.73457 | 0.90638 | 0.68264 | 0.89151 |
| ConvLSTM | 0.95729 | 0.94636 | 0.95220 | 0.94900 | 0.94636 | 0.98133 | 0.95220 | 0.97995 |
| **48 hour prediction Model** | **Accuracy** | **Precision** | **Recall** | **F1** | **PPV** | **NPV** | **Sensitivity** | **Specificity** |
| SVM | 0.81682 | 0.91825 | 0.73365 | 0.79571 | 0.91825 | 0.94395 | 0.73365 | 0.91261 |
| LSTM | 0.79510 | 0.81192 | 0.72112 | 0.75577 | 0.81192 | 0.92440 | 0.72112 | 0.91546 |
| ConvLSTM | 0.92177 | 0.91497 | 0.90908 | 0.90947 | 0.91497 | 0.97190 | 0.90908 | 0.96827 |
| Abbreviation: PPV: Positive Predictive Value; NPV: Negative predictive value. | | | | | | | | |

In terms of ConvLSTM models, the AUC of 24-hour prediction AKI is 99.78%, 48-hours AKI 99.43% in test datasets. The F1 score is an indicator that combines precision and recall. It is mostly used for the measurement of categorical variables. The F1 score predicting the 24-hour AKI model and predicting the 48-hour AKI model are 94.90% and 90.95% respectively. External validations on the eICU middle-aged and older patients' cohort suggest that our model has the potential to be promoted and applied in other hospitals. The results demonstrate that the models are the best performance in the area of predicting weather middle-aged and older patients in ICUs will develop acute kidney injury.

# Discussion

Here we present the three models, efficient computational models for daily prediction of the development of AKI in critically ill middle-aged and older patients. The ConvLSTM models surpassed any previous model to predict whether older patients have AKI or not with solid evaluation metrics. Traditional statistical prediction models in medicine have made a great contribution to find related variables about the current treatment decision support. However, they have some limitations that could rarely be timeliness and accuracy. The proposed deep learning model was trained as an effort to overcome the limitations presented by traditional strategies of building prediction models. Through developing the deep neural networks, we demonstrate that deep learning can handle lots of variables which may be predictors. In addition, our results indicate that leveraging clinical data of time series as well as a deep learning model helps to efficiently learn whether AKI development in a specific group of people and boost the prediction performance. The ConvLSTM model using the daily laboratory examination and vitals can be used to predict real-time dynamics at different time points, facilitating decision-making for the physician in ICU throughout the patient's entire hospitalization.

For over a decade, the study of biomarkers and urine volumes that can recognize AKI early and more reliably than serum creatinine elevated has been the focus of research[23–26] The machine learning model improves the accuracy of predicting AKI stages, which may be of great significance for middle-aged and older patients treated in ICU, especially for the patients with severe AKI in the future. For example, the model predicts that the patient would develop to AKI stage3 after 48 hours, based on which doctors in ICU could intervene in advance, such as disabling kidney damage drugs, rehydration therapy, or early CRRT. If the patient is identified stage 1 AKI by the model, then physicians can choose rehydration therapy and careful use of kidney damage drugs to protect kidney organs. About the application, the data input into these models is based on daily routine laboratory inspection and basic information, which can promote in most hospitals because the models don't need the latest biological indicators.

The electronic health record (EHR) makes it possible to obtain more records, and the development of modern more powerful computing devices makes it possible to use machine learning-based models to calculate whether patients have diseases in the future. For example,

the simple real-time model launched by Yale School of Medicine trained discrete-time logistic-based machine learning models on clinical data from EHR and deployed the trained model as an embedded in EHR system service that can absorb real-time information from the hospitalized patient to provide a timely prediction of the probability of patient acquiring AKI[27]. Similarly, the deep learning model we developed the performance is better than previous models and associated with the EHR system database, the model may provide hourly AKI stage assessment in the specific group for the next 24 and 48 hours.

Despite the accuracy advantages and potential clinical implications of deep learning demonstrated in these analyses, there are certain obstacles to the application of computer models in the clinic settings, such as doctors do not believe and do not understand algorithms. First, a most of clinicians are more familiar with linear regression and logistical regression models than deep learning techniques, and logistic modeling shows the regression coefficient can be converted into OR value, which helps doctors to understand the strength of the relationship between variables and the outcome of the patient[28, 29]. Despite deep neural network is excellent in building disease prediction models, the correlation between each variable and the outcome is poorly explained, so the deep learning algorithm is not good at giving suggestions of treatment advice based on the prediction. Reinforcement learning may be more suitable for giving better decisions in the clinical environment, in which we will study the application of reinforcement learning in clinical settings.

There are several limitations to this study. Firstly, it was a retrospective study, we did not combine with the EHR to automatically import data to our models. With the availability of time series data of examination items of patients and importing those data into models directly, fully automated prediction at different time points becomes possible and is an avenue for future work. Secondly, after making the prediction, the ConvLSTM model does not give the doctors what kind of decision should be made. In future work, models can be built to assist doctors in decision-making to help them rather than replace them. Thirdly, this study is not using all data during hospitalization, so as treatment and medication improve, the change of probability of mortality may not show up. This can be improved by collected more old patients with AKI.

# Conclusion

In conclusion, this study developed an efficient model to predict the real-time AKI stage after 24 hours and 48 hours of middle-aged and elderly patients in ICU by deep learning based on large EHR data. It is the first to

demonstrate the strong ability of the deep neural network in this area, and the prediction model we present is the best performance than any previous. With the prevalence of computing power and integrated EHRs at the bedside, predictive models developed with these techniques have the potential to be incorporated into routine patient care as well as to support medical care quality improvement.

# Declarations

**Ethics approval and consent to participate:**

Not applicable

**Consent for publication:**

Not applicable

**Availability of data and material:**

All of the data that support the findings of this study are available from MIMIC III & eICU datasets. All models were developed and applied using Tensorflow 1.7. Detailed parameters are given in the method section, and the neural network code is public in the Github (https://github.com/ZEROICEWANG/Older-aki). The code for data extraction is available on the respective guidance sites.

**Competing interests:**

The authors declare that they have no competing interests.

**Funding:**

This work was supported by the Special fund for clinical research of Wu Jieping Medical Foundation (URL: http://www.wjpmf.org/en/; No.HRJJ20180722).

**Authors' contributions:**

Fuxing Deng analyzed the data and wrote the paper. Fuxing Deng and Hui Wang built those models. Buyao Zhang reviewed the data analyzation analyzed the data. Shuangping Zhao designed the study.

**Acknowledgments:**

In writing this paper, I have benefited from Leo Anthony Celi, Zhang Zhongheng, and the Datathon in Beijing, Nov 2019. Leo's speech gave me a more comprehensive understanding of the database. Zhang's lectures inspired me to data exploration. I acquired information and inspiration from the Datathon and thanks to the organizers.

# References

1. Lameire NH, Bagga A, Cruz D, De Maeseneer J, Endre Z, Kellum JA, Liu KD, Mehta RL, Pannu N, Van Biesen W *et al*: **Acute kidney injury: an increasing global concern**. *The Lancet* 2013, **382**(9887):170-179.
2. Mehta RL, Burdmann EA, Cerdá J, Feehally J, Finkelstein F, García-García G, Godin M, Jha V, Lameire NH, Levin NW *et al*: **Recognition and management of acute kidney injury in the International Society of Nephrology**

**0by25 Global Snapshot: a multinational cross-sectional study**. *The Lancet* 2016, **387**(10032):2017-2025.

3. Yang L, Xing G, Wang L, Wu Y, Li S, Xu G, He Q, Chen J, Chen M, Liu X *et al*: **Acute kidney injury in China: a cross-sectional survey**. *The Lancet* 2015, **386**(10002):1465-1471.
4. Xue JL, Daniels F, Star RA, Kimmel PL, Eggers PW, Molitoris BA, Himmelfarb J, Collins AJ: **Incidence and mortality of acute renal failure in Medicare beneficiaries, 1992 to 2001**. *J Am Soc Nephrol* 2006, **17**(4):1135-1142.
5. Kane-Gill SL, Sileanu FE, Murugan R, Trietley GS, Handler SM, Kellum JA: **Risk factors for acute kidney injury in older adults with critical illness: a retrospective cohort study**. *Am J Kidney Dis* 2015, **65**(6):860-869.
6. Schmitt R, Coca S, Kanbay M, Tinetti ME, Cantley LG, Parikh CR: **Recovery of Kidney Function After Acute Kidney Injury in the Elderly: A Systematic Review and Meta-analysis**. *American Journal of Kidney Diseases* 2008, **52**(2):262-271.
7. Anderson S, Eldadah B, Halter JB, Hazzard WR, Himmelfarb J, Horne FM, Kimmel PL, Molitoris BA, Murthy M, O'Hare AM *et al*: **Acute kidney injury in older adults**. *J Am Soc Nephrol* 2011, **22**(1):28-38.
8. Ferenbach DA, Bonventre JV: **Mechanisms of maladaptive repair after AKI leading to accelerated kidney ageing and CKD**. *Nature Reviews Nephrology* 2015, **11**(5):264-276.
9. He L, Wei Q, Liu J, Yi M, Liu Y, Liu H, Sun L, Peng Y, Liu F, Venkatachalam MA *et al*: **AKI on CKD: heightened injury, suppressed repair, and the underlying mechanisms**. *Kidney International* 2017, **92**(5):1071-1083.
10. Ishani A, Xue JL, Himmelfarb J, Eggers PW, Kimmel PL, Molitoris BA, Collins AJ: **Acute kidney injury increases risk of ESRD among elderly**. *J Am Soc Nephrol* 2009, **20**(1):223-228.
11. Schmitt R, Cantley LG: **The impact of aging on kidney repair**. *Am J Physiol Renal Physiol* 2008, **294**(6):F1265-1272.
12. Uchino S, Kellum JA, Bellomo R, Doig GS, Morimatsu H, Morgera S, Schetz M, Tan I, Bouman C, Macedo E *et al*: **Acute Renal Failure in Critically Ill PatientsA Multinational, Multicenter Study**. *JAMA* 2005, **294**(7):813-818.
13. Xu X, Nie S, Liu Z, Chen C, Xu G, Zha Y, Qian J, Liu B, Han S, Xu A *et al*: **Epidemiology and Clinical Correlates of AKI in Chinese Hospitalized Adults**. *Clin J Am Soc Nephrol* 2015, **10**(9):1510-1518.
14. Makridakis S: **The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms**. *Futures* 2017, **90**:46-60.
15. Struble TJ, Alvarez JC, Brown SP, Chytil M, Cisar J, DesJarlais RL, Engkvist O, Frank SA, Greve DR, Griffin DJ *et al*: **Current and Future Roles of Artificial Intelligence in Medicinal Chemistry Synthesis**. *J Med Chem* 2020.
16. Chuai G, Ma H, Yan J, Chen M, Hong N, Xue D, Zhou C, Zhu C, Chen K, Duan B *et al*: **DeepCRISPR: optimized CRISPR guide RNA design by deep learning**. *Genome Biol* 2018, **19**(1):80.
17. Chang K, Bai HX, Zhou H, Su C, Bi WL, Agbodza E, Kavouridis VK, Senders JT, Boaro A, Beers A *et al*: **Residual Convolutional Neural Network for the Determination of IDH Status in Low- and High-Grade Gliomas from MR Imaging**. *Clin Cancer Res* 2018, **24**(5):1073-1081.
18. Rashidi HH, Sen S, Palmieri TL, Blackmon T, Wajda J, Tran NK: **Early Recognition of Burn- and Trauma-Related Acute Kidney Injury: A Pilot Comparison of Machine Learning Techniques**. *Sci Rep* 2020, **10**(1):205.
19. Huang C, Murugiah K, Mahajan S, Li SX, Dhruva SS, Haimovich JS, Wang Y, Schulz WL, Testani JM, Wilson FP *et al*: **Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: A retrospective cohort study**. *PLoS Med* 2018, **15**(11):e1002703.

20. Deng F, Peng M, Li J, Chen Y, Zhang B, Zhao S: **Nomogram to predict the risk of septic acute kidney injury in the first 24 h of admission: an analysis of intensive care unit data**. *Ren Fail* 2020, **42**(1):428-436.
21. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG: **MIMIC-III, a freely accessible critical care database**. *Sci Data* 2016, **3**:160035.
22. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O: **The eICU Collaborative Research Database, a freely available multi-center database for critical care research**. *Sci Data* 2018, **5**:180178.
23. Marx D, Metzger J, Pejchinovski M, Gil RB, Frantzi M, Latosinska A, Belczacka I, Heinzmann SS, Husi H, Zoidakis J *et al*: **Proteomics and Metabolomics for AKI Diagnosis**. *Seminars in Nephrology* 2018, **38**(1):63-87.
24. Alge JL, Arthur JM: **Biomarkers of AKI: a review of mechanistic relevance and potential therapeutic implications**. *Clin J Am Soc Nephrol* 2015, **10**(1):147-155.
25. Kellum JA, Sileanu FE, Murugan R, Lucko N, Shaw AD, Clermont G: **Classifying AKI by Urine Output versus Serum Creatinine Level**. *J Am Soc Nephrol* 2015, **26**(9):2231-2238.
26. Zhang P, Yi L, Qu S, Dai J, Li X, Liu B, Li H, Ai K, Zheng P, Qiu S *et al*: **The Biomarker TCONS_00016233 Drives Septic AKI by Targeting the miR-22-3p/AIFM1 Signaling Axis**. *Mol Ther Nucleic Acids* 2020, **19**:1027-1042.
27. Simonov M, Ugwuowo U, Moreira E, Yamamoto Y, Biswas A, Martin M, Testani J, Wilson FP: **A simple real-time model for predicting acute kidney injury in hospitalized patients in the US: A descriptive modeling study**. *PLoS Med* 2019, **16**(7):e1002861.
28. Liu X, Ye Y, Mi Q, Huang W, He T, Huang P, Xu N, Wu Q, Wang A, Li Y *et al*: **A Predictive Model for Assessing Surgery-Related Acute Kidney Injury Risk in Hypertensive Patients: A Retrospective Cohort Study**. *PLoS One* 2016, **11**(11):e0165280.
29. Luo M, Yang Y, Xu J, Cheng W, Li XW, Tang MM, Liu H, Liu FY, Duan SB: **A new scoring model for the prediction of mortality in patients with acute kidney injury**. *Sci Rep* 2017, **7**(1):7862.

# Figures

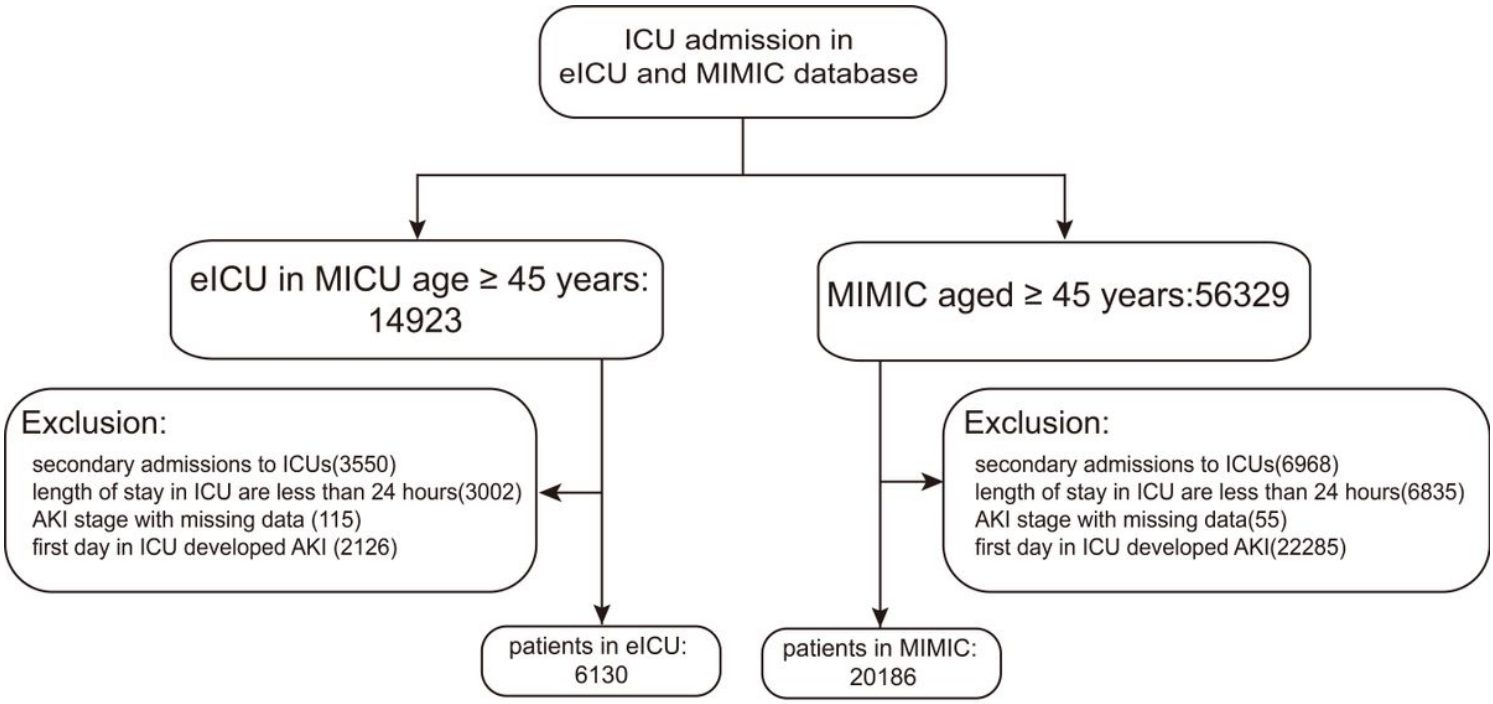


Fig 1

**Figure 1**

the patient selection and reasons for exclusion.

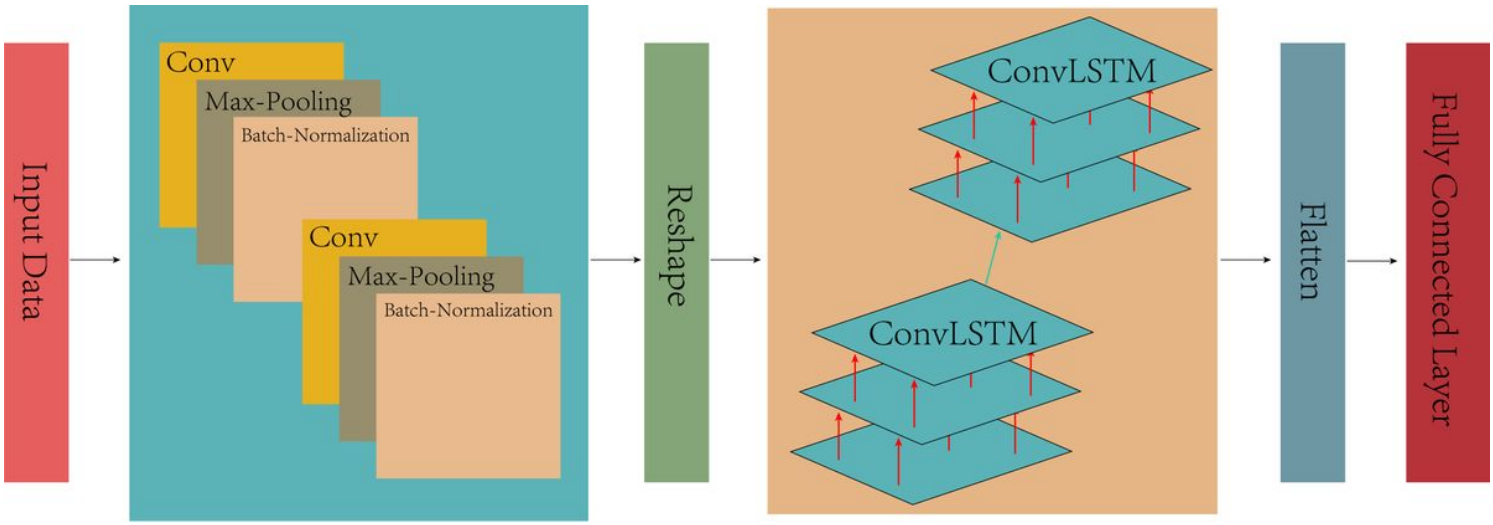


Fig. 2

**Figure 2**

The schematic diagram of the model to predict AKI stage for patients;

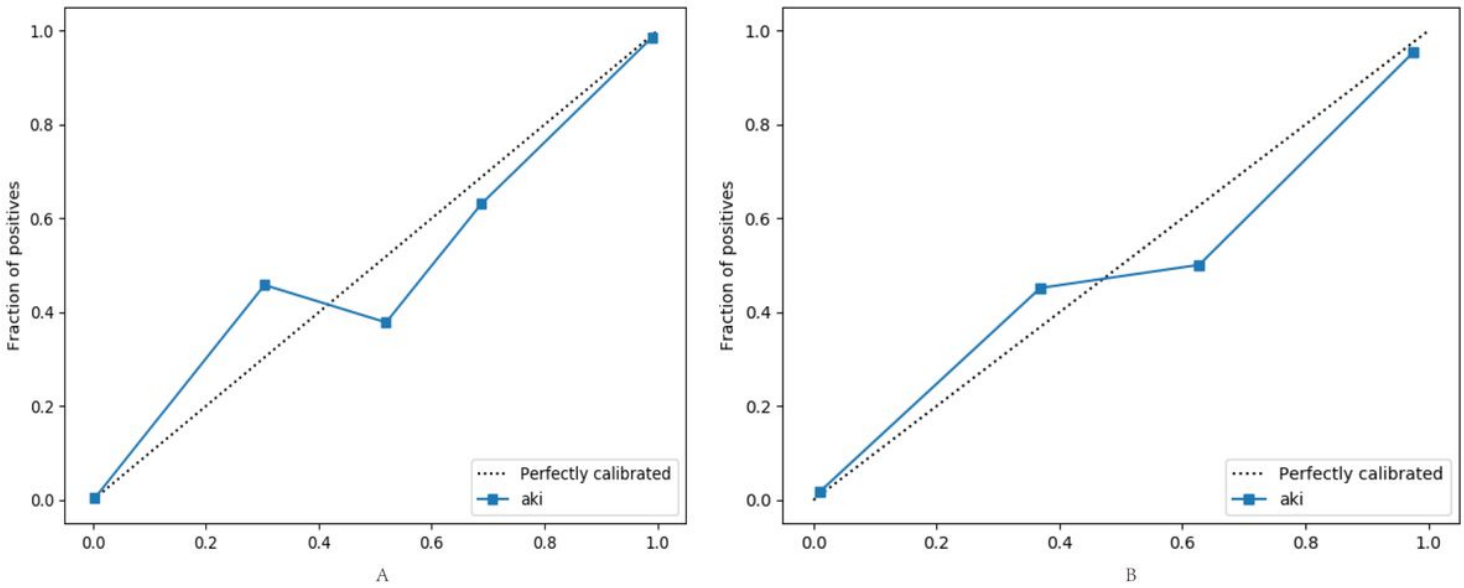


Figure 3

**Figure 3**

3A: The calibration curve of prediction of after 24-hour AKI stage; 3B: The calibration curve of prediction of after 48-hour AKI stage.

## Supplementary Files

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

- missingMIMIC.pdf
- missingMIMIC.pdf
- missingeICU.pdf
- missingeICU.pdf
- AUC.docx
- AUC.docx