BMJ Open

# Prospective Validation for Severe Acute Kidney Injury Predicting Model based on transcontinental databases



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2021-054092 |
| Article Type: | Original research |
| Date Submitted by the Author: | 16-Jun-2021 |
| Complete List of Authors: | Liang, Qiqiang; Zhejiang University School of Medicine Second Affiliated Hospital<br>Xu, Yongfeng; HealSci Technology Co<br>Zhou, Yu; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Xinyi; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Juan; Zhejiang University School of Medicine Second Affiliated Hospital<br>Huang, Man; Zhejiang University School of Medicine Second Affiliated Hospital, |
| Keywords: | Acute renal failure < NEPHROLOGY, INTENSIVE & CRITICAL CARE, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Prospective Validation for Severe Acute Kidney Injury Predicting Model based on transcontinental databases

Qiqiang Liang[1]; Yongfeng Xu[2]; Yu Zhou[1]; Xinyi Chen[1], MM; Juan Chen[1]; Man Huang[1*]

[1]General Intensive Care Unit, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

[2]BSc, HealSci Technology Co., Ltd, Beijing China.

*Corresponding author: Man Huang, MD PhD.

Department: General Intensive Care Unit

Hospital: Second Affiliated Hospital of Zhejiang University School of Medicine

Address: No. 1511, Jianghong Road, Bingjiang District, Hangzhou City, Zhejiang Province

Tel: + 86 571 89713427

Fax: + 86 571 89713427

E-mail: huangman@zju.edu.cn.

ORCID: 0000-0002-0842-2254

**Abstract:**

Objectives: There are many studies of acute kidney injury (AKI) diagnosis models lack of external validation and prospective validation. We constructed the models using three databases to predict severe AKI within 48 hours in intensive care unit (ICU) patients.

Design: A retrospective and prospective cohort study.

Setting: We studied critically ill patients in our database (SHZJU-ICU) and two other public databases, the Medical Information Mart for Intensive Care (MIMIC) and AmsterdamUMC databases, including basic demographics, vital signs, and laboratory results. We predicted the diagnosis of severe AKI in patients in the next 48 hours using machine-learning algorithms with the three databases. Then, we carried out real-time severe AKI prediction in the prospective validation study at our centre for one year.

Participants: All patients included in three databases with uniform exclusion criteria.

Primary and secondary outcome measures: Effect evaluation index of prediction models.

Results: We included 58492 patients, and a total of 5257 (9.0%) patients met the definition of severe AKI. In the internal validation of the SHZJU-ICU and MIMIC databases, the best area under the receiver operating characteristic curve (AUROC) of the model was 0.86. The external validation results by AmsterdamUMC database were also satisfactory, with the best AUROC of 0.86. A total of 2532 patients were admitted to the center for prospective validation; 358 positive results were predicted, and 344 patients were diagnosed with severe AKI, with an accuracy of 83.5% and an AUROC of 0.84.

Conclusion: The prediction model of severe AKI exhibits promises as a clinical application based on dynamic vital signs and laboratory results of multi-center databases with prospective and external validation.

**Keywords**: Machine learning; Acute kidney injury; Real-time prospective validation; External

validation.

**Strengths and limitations of this study**:

An important verification step from artificial intelligence research to clinical use;

Three large database containing different national populations and regions;

The prediction model with excellent performance in complete data validation;

Differences in the samples proportion of three databases;

The dimensions of variables are not rich enough.

**Introduction:**

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalization and mortality[1]. AKI is divided into three types according to the various aetiologies: prerenal (renal hypoperfusion), intrarenal (vascular, glomerular, or tubulointerstitial lesions), and postrenal (urinary tract obstruction)[2]. Although nearly all diseases associated with ICU admission may cause AKI, acute tubular necrosis (ATN) and prerenal azotaemia are the most common causes[3].

All AKI diagnostic criteria including the latest Kidney Disease: Improving Global Outcomes (KDIGO) standard are currently based on the creatinine level and urine volume[4]. However, the increase in the creatinine level or decrease in the urine volume lags the onset of AKI[2]. Many studies have suggested that early diagnosis and treatment of reversible AKI can reduce mortality[5]. Therefore, the creatinine level and urine volume are not satisfactory to meet clinical diagnostic demands. Consequently, many researchers have tried to develop an early warning model by analysing the risk factors for AKI[6].

Patient complications, such as diabetes, hypertension, cardiovascular disease, chronic liver disease,

sepsis, and trauma, are identified as important risk factors for AKI[7]. The AKI prediction model and scoring system developed based on high-risk factors has gradually become the focus of research considering the lower clinical application threshold compared with that of new biomarkers[6]. Although most previous prediction models use the multiple logistic regression model, a variety of AKI prediction models based on machine learning have resulted in satisfactory outcomes[6]. Since the first AKI prediction model study based on artificial intelligence was published in 2016, researchers have built more than 20 published AKI prediction models successively by using local or multi-centre databases[6,8-14]. The results indicate that these models can predict the occurrence of AKI and the need for renal replacement therapy (RRT) within 24 or 48 hours, with accuracies ranging from 81% to 97%[6,15]. In addition, many studies have focused on subspecialized conditions, including cardiac surgery, trauma, and burns[14-16]. However, the common defect in these studies is the lack of external validation and prospective validation, which causes the prediction model to deviate from the clinical scenarios and limits extrapolation beyond the scope of the data.

In this study, we built models to predict AKI within 48 hours in critically ill patients by using transcontinental three databases. Then, we evaluated the clinical effect of the model through a one-year prospective validation at our centre.

## Materials and Methods:

### Study Population

We collected patients using three ICU databases and prospectively validated the models in our centre. The first database was our centre general ICU database (SHZJU-ICU) of the Second Affiliated Hospital of Zhejiang University School of Medicine, an academic teaching hospital. Since its establishment in

2017, it has included 12000 ICU patients' data and is updated daily. The Medical Information Mart for Intensive Care (MIMIC) III database, the second one, is an open ICU database provided by the Massachusetts Institute of Technology and includes nearly 60000 ICU patients from North America[17]. Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23000 patients admitted to ICUs in parts of Europe[18]. The research flow chart is shown in Figure 1. The study was evaluated and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine as study number 2019-078.

**Study definition**

In this study, the diagnosis of AKI was confirmed based on three stages according to the KDIGO criteria[4]. We defined the patients who met the KDIGO AKI II and III criteria as severe AKI groups and the others as negative groups. We excluded patients with lack of creatinine measurements during admission, patients with creatinine baseline more than 3.0 mg/dL at admission, patients who met severe AKI diagnosis within 24 hours, and patients who used RRT within 48 hours after admission[19]. In addition, we excluded pregnant women, patients younger than 14 years old, and patients hospitalized in the ICU for fewer than 3 days. After the patient was admitted to ICU, we performed a prediction every 24 hours and recorded a prediction time. If the patient was diagnosed with severe AKI within 48 hours, the predictive time was defined as a positive predictive point, and the others were defined as a negative point.

**Data Collection**

The variables included demographic data, vital signs, basic and primary diseases, laboratory results, important operation records, and drug records. Comorbidity included hypertension, diabetes, cardiopathy, liver disease, and malignant tumors. The primary disease was the main cause of admission to the ICU

following the ICD-10 codes. The vital signs and clinical laboratory results were transformed into different variables according to the average, variance, maximum, minimum, and final value before diagnosis. We used the Pearson correlation coefficient to analyze the characteristics of the variables, selected the variables with high correlation. We transformed the MIMIC and AmsterdamUMC databases according to our centre database structure, unifying the unit and diagnostic codes. We deleted more than 50% of the missing values and replaced the remaining missing values with multiple interpolations. All Missing data between three databases and values included in the model shown in supplementation file Table S1.

**Model construction and external validation**

The ratio of the training and internal validation sets was 4:1. The two database training sets were mixed into a new training set. There were more negative data than positive data, so we randomly sampled the negative datasets and constructed a new data subset with a sampling ratio of positive and negative data of 1:5. We used multiple logistic regression, random forest, XGBoost, AdaBoost, LightGBoost, gradient boosting decision tree (GBDT), and debug to assess the variables and model-related parameters by the fivefold cross-validation method. After the models were built, we used the SHZJU-ICU and MIMIC test sets for internal validation of the model and the AmsterdamUMC database for external validation. The most appropriate cut-off value was determined according to the K-S curve. The prediction model represents the results of each prediction with a probability between 0 to 1.0. We define results more than 0.6 as high-risk, that is, positive results, and the rest as negative results. Through internal validation and external validation, we calibrated the model by adjusting the super-parameters and using the Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. The model was visualized and analyzed by a visual programming scheme. All model building and validation

processes were performed in Python 3.6.

**Prospective validation**

The prospective research period was 2020.01.01 to 2020.12.31. We collected real-time data when patients were admitted to the ICU, transformed the data according to the requirements, and formed a complete sample for the prediction model after passing the integrity test. We had established a visualization scheme and allowed researchers to review the predictions daily. The daily prediction results were not publicly accessible during the study to avoid affecting clinicians' decisions, but the diagnosis results were available to the researchers as visual graphics. We sampled the 20% predicted data every month and deleted samples with more than 50% missing values to ensure data correctness. The criteria to terminate prediction were A. a positive diagnosis; B. Transfer out of ICU or death with negative diagnosis. All diagnosis of severe AKI needs to be reviewed by two ICU attending physicians independently, and if the they have different opinion, the third one will be appealed.

**Statistical analysis:**

The population characteristics were reported as the medians and inter-quartile ranges (IQRs) for skewed data and the means and standard deviations for normally distributed data. The independent sample T-test was used for normally distributed data, and the rank-sum test was used for the rest. Dichotomous variables were assessed by the chi-square test, and a P value less than 0.05 was considered statistically significant. The non-normally distributed data were analyzed by exponential transformation and logarithmic transformation, and the variables that did not reach a normal distribution were deleted. The effect of the model was evaluated by parameters such as the AUROC, accuracy, specificity, and F1-score.

**Results:**

According to the inclusion criteria and exclusion criteria, we selected 58492 patients from three databases who met the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36690 patients from the MIMIC database, and 15341 patients from the AmsterdamUMC database. A total of 5257 (9.0%) patients met the definition of severe AKI (11.8% in SHZJU-ICU, 7.6% in MIMIC, and 10.9% in AmsterdamUMC). The distributions of age and sex in the three centers were similar, but the differences in race were large. Asian patients accounted for more than 99% in the SHZJU-ICU database, and only approximately 2.5% in the MIMIC database. White people accounted for more than 70% of the MIMIC database. In addition, patients from the MIMIC database had a higher incidence of the tumor, liver cirrhosis, diabetes, and hypertension. Patients in the AmsterdamUMC and SHZJU-ICU databases had a higher proportion of mechanical ventilation and overall survival rate. Severe AKI patients had longer ICU hospital stays and higher mortality. More details are presented in Table 1.

Table 1: Clinical Demographics and Outcomes in patients with or without severe acute kidney injury.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | MIMIC Negative (N=33879) | MIMIC Severe AKI (N=2811) | P-value | AmsterdamUMC Negative (N=13661) | AmsterdamUMC severe AKI (N=1680) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 62.0 [45.0-76.0] | 64.0 [52.0-76.0] | <0.001 | 64.5 [54.5,74.5] | 64.3 [52.5,78.5] | 0.04 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 18976 (56.0) | 1524 (54.2) | 0.06 | 8920 (65.3) | 1054 (63.9) | 0.03 |
| **Race** | | | | | | | / | / | / |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 23796 (70.2) | 2005 (71.3) | 0.23 | / | / | / |
| **Black** | 1 (0.02) | 0 | 1 | 2455 (7.2) | 199 (7.1) | 0.45 | / | / | / |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 918 (2.7) | 70 (2.5) | 0.54 | / | / | / |
| **Other** | 0 | 1 (0.1) | 1 | 6710 (19.8) | 537 (19.1) | 0.34 | / | / | / |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 1.1 (0.41) | 1.1 (0.50) | 0.04 | 1.2 (0.35) | 1.1 (0.38) | <0.001 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (12) | 23 (12) | <0.001 | 18 (13) | 23 (10) | <0.001 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 6652 (19.6) | 752 (26.8) | <0.001 | / | / | / |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 1180 (3.5) | 426 (15.2) | <0.001 | / | / | / |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 4840 (14.3) | 336 (12.0) | 0.001 | / | / | / |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 7250 (21.4) | 789 (28.1) | <0.001 | / | / | / |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 16328 (48.2) | 1513 (53.8) | <0.001 | / | / | / |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 13009 (38.4) | 1503(53.5) | <0.001 | 7718(56.5) | 1248(74.3) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 12455 (36.7) | 1430 (50.8) | <0.001 | / | / | / |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 51.0 [28.5-110.3] | 103.4 [50.1-261.3] | <0.001 | 24.0 [19.8,63.5] | 142.0 [45.4,394.2] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 25648 (75.7) | 1356 (48.2) | <0.001 | 10942(80.1) | 1033 (61.5) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

There were significant differences in the important parameters of the variables among the different models (see Figure S1). However, the trend of the creatinine level in the past week was still an important variable, followed by urine volume, blood urea nitrogen level, temperature, and length of ICU stay. The cut-off value used to distinguish between a negative and positive prediction was determined by the K-S curve, with value of 0.423 (see Figure S2). The GBDT model had the best prediction effect in the test set, followed by XGBoost and LightGBoost. In the two central internal validation sets, the two best-performing machine learning algorithms with great AUROC are LightGBoost (SHZJU-ICU of 83.2%, MIMIC of 86.0%) and XGBoost (SHZJU-ICU 85.9%, MIMIC 85.6%), as detailed in Figure 2. Overall, the sensitivity (SHZJU-ICU 0.84, MIMIC 0.83) and the negative predictive value (SHZJU-ICU 0.90, MIMIC 0.90) of the predictive model were high, but the specificity was general (SHZJU-ICU 0.79,

MIMIC 0.75), as shown in Table 2. In the external validation based on AmsterdamUMC database, the overall model validation effect results were satisfactory, and XGBoost had the best performance, with an AUROC of 0.84, as shown in Figure 2 and Table 2. We built a visual prediction interface based on the prediction model (supplementation file Figure S3).

Table 2: Model validation results by three databases with machine learning algorithm

| Model | AUROC | Accuracy | Sensitivity | Sepcificity | PPV | NPV | F1 |
|---|---|---|---|---|---|---|---|
| ***Internal validation with SHZJU-ICU database*** | | | | | | | |
| ***Logistic regression*** | *0.748* | *0.662* | *0.834* | *0.576* | *0.496* | *0.874* | *0.622* |
| ***LightGBoost*** | *0.832* | *0.741* | *0.839* | *0.692* | *0.576* | *0.896* | *0.683* |
| ***GBDT*** | *0.845* | *0.765* | *0.843* | *0.725* | *0.606* | *0.902* | *0.705* |
| ***AdaBoost*** | *0.806* | *0.721* | *0.824* | *0.67* | *0.555* | *0.884* | *0.663* |
| ***Random Forest*** | *0.821* | *0.763* | *0.71* | *0.789* | *0.627* | *0.845* | *0.666* |
| ***XGBoost*** | ***0.859*** | *0.779* | *0.81* | *0.763* | *0.631* | *0.889* | *0.709* |
| ***Internal validation with MIMIC database*** | | | | | | | |
| ***Logistic regression*** | *0.733* | *0.695* | *0.643* | *0.72* | *0.535* | *0.801* | *0.584* |
| ***LightGBoost*** | ***0.86*** | *0.768* | *0.822* | *0.741* | *0.613* | *0.893* | *0.702* |
| ***GBDT*** | *0.846* | *0.765* | *0.786* | *0.755* | *0.616* | *0.876* | *0.691* |
| ***AdaBoost*** | *0.837* | *0.732* | *0.831* | *0.683* | *0.567* | *0.89* | *0.674* |
| ***Random Forest*** | *0.832* | *0.738* | *0.791* | *0.712* | *0.578* | *0.872* | *0.668* |
| ***XGBoost*** | *0.856* | *0.758* | *0.833* | *0.721* | *0.598* | *0.895* | *0.695* |
| ***External validation with AmsterdamUMC database*** | | | | | | | |
| ***Logistic regression*** | *0.704* | *0.767* | *0.516* | *0.893* | *0.706* | *0.787* | *0.596* |
| ***LightGBoost*** | *0.859* | *0.763* | *0.827* | *0.731* | *0.606* | *0.894* | *0.7* |
| ***GBDT*** | *0.861* | *0.764* | *0.84* | *0.727* | *0.606* | *0.901* | *0.704* |
| ***AdaBoost*** | *0.85* | *0.755* | *0.813* | *0.726* | *0.597* | *0.886* | *0.689* |
| ***Random Forest*** | *0.82* | *0.743* | *0.77* | *0.729* | *0.587* | *0.864* | *0.666* |
| ***XGBoost*** | ***0.865*** | *0.75* | *0.873* | *0.688* | *0.584* | *0.916* | *0.7* |
| ***Prospective validation with SHZJU-ICU*** | | | | | | | |
| ***Logistic regression*** | *0.758* | *0.772* | *0.648* | *0.834* | *0.662* | *0.826* | *0.655* |
| ***LightGBoost*** | *0.819* | *0.796* | *0.596* | *0.895* | *0.74* | *0.816* | *0.66* |
| ***GBDT*** | *0.827* | *0.781* | *0.706* | *0.818* | *0.66* | *0.848* | *0.683* |
| ***AdaBoost*** | *0.808* | *0.766* | *0.686* | *0.805* | *0.638* | *0.837* | *0.661* |
| ***Random Forest*** | *0.804* | *0.755* | *0.715* | *0.775* | *0.613* | *0.845* | *0.66* |
| ***XGBoost*** | ***0.841*** | *0.779* | *0.724* | *0.807* | *0.652* | *0.854* | *0.686* |

*AUROC*: Area Under Receiver Operating Characteristic; *PPV*: Positive Predictive Value; *NPV*: Negative Predictive Value; *SHZJU-ICU*: The general ICU database of the second affiliated

hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *GBDT*: Gradient Boosted Decision Tree.

According to the inclusion and exclusion criteria, we delete 267 patients among 94 patients with creatinine baseline more than 3.0 mg/dL at admission, 39 patients met severe AKI diagnosis within 24 hours, and 26 patients who used RRT within 48 hours after admission, 108 patients hospitalized in the ICU for fewer than 3 days. A total of 2532 patients were admitted to our centre for prospective validation, and the prediction model made 16858 times predictions. In the prospective cohort, there was no significant difference in age, gender, baseline creatinine and urea nitrogen, and complications. The proportion of mechanical ventilation and the ICU stay time in AKI patients were longer with higher mortality. Above all, there was no significant difference between the prospective and the retrospective cohort. More detail sees in supplementation file Table S2. In the end, 358 positive results were predicted, and the rest were negative results. There are 344 patients with severe AKI were diagnosed and the prediction accuracy was 83.5%. The model with the highest area under the curve was XGBoost, 0.84. The results of the prospective study are similar to those of the external validation of the model, and are relatively stable. More detail is presented in Figure 3 and Table 2.

**Discussion:**

In this study, we built predictive models by machine learning to predict the incidence of severe AKI with three databases in different regions and in the next 48 hours. After internal and external validation, prospective validation over one year was carried out to verify the model effects. The three databases come from three countries that are in Asia, Europe, and North America, which proves that the model is universal to some extent.

Despite the huge amount of data, many databases are still not suitable for prospective research

because they are not updated promptly. Tomasev, N and colleagues have provided research on AKI prediction models with a large amount of data[10]. The study covered 703782 adult patients with 6 billion individual items, including 620000 elements. In this study, a depth neural network model was used for real-time prediction. A total of 55.8% of severe AKI patients were predicted within the first 48 hours, although each accurate prediction was accompanied by two mispredictions[10]. This study provided a new scheme for real-time prediction and indicated that we should prospectively evaluate and independently validate models to explore their effectiveness. In a prospective study, Marine, F compared an AKI prediction model with clinicians in 252 patients and found that the clinical effect of the random forest model for predicting AKI-II/III was equivalent to that of clinicians. Our prediction model graphical visualization of the model was installed in the centre's database for better usage. In addition, our database is updated daily to achieve daily predictions and present the results to researchers. In the prospective validation of our study, the stability of the prediction model confirmed its promise, which provides a basis for future research.

There are many studies of artificial intelligence for predicting the occurrence of AKI, but most of them are single-centre studies, and the extrapolation effect has been controversial. In 2016, Koyner, J, and his colleague published the first study of an AKI prediction model based on multi-centre data. In all 202961 patients, 17541 (8.6%) had AKI, 4251 (3.5%) had AKI-II, and 1242 (0.6%) had AKI-III. A multivariate logistic regression model was used to predict AKI in this study with an AUROC of 0.74 (95% CI=0.74). With the classification of AKI, the AUROC of the prediction model gradually increased to 0.84[8]. Subsequently, the study team used a new machine learning algorithm, to build a more accurate model to predict the occurrence of AKI-II, with an AUROC of 0.9 within 24 hours and 0.87 within 48 hours[19]. Recently, the research team included data from two other centres, namely, LUMC (N=200613) and NUS

(N=246895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85 to 0.86, suggesting that the artificial intelligence model has stable predictive ability[12]. This series of studies included many data points, suggesting the feasibility of artificial intelligence in the diagnosis of AKI, but the proportion of positive patients (3.5%) and ICU patients (30%) was too low to properly predict AKI. Our research is similar to the above. The SAHZJU-ICU database is a single-centre database representing south-eastern China, and the MIMIC database is a well-known open ICU database in the United States. The AmsterdamUMC database is a public database located in Europe. The population structure and diseases in the three databases are complete but different in the distribution of complications and race. Therefore, it provides a prediction model with unparalleled stability compared with other studies.

Limitations:

This retrospective multi-centre study was unable to carry out more clinical feature mining and comparison because of different data structures. The differences between the three databases partly reflect some demographic differences between Europe, the United States and China, resulting in a decline in the accuracy of the prediction model. There are some differences in the number of patients included in the three databases, which may affect the choice of variables. Second, in the prospective data study in 2020, there may be deviations in the inclusion of patients in the centre, thus affecting the interpretation of the follow-up forward-looking results. Finally, given the low incidence of severe AKI and the great difference in the proportion of positive and negative samples, the data may be accidental.

**Conclusion:**

Based on databases of patients of different races from different countries, we constructed stable

machine learning models to predict the occurrence of AKI in the next 48 hours. Prospective validation through the implementation of an updated local database is an effective exploration of further research.

Data Availability Statement:

The model code can be obtained by email if readers need it, but we cannot guarantee that all the code will be provided. Two public databases can be applied from the official website of their respective databases.

Ethics approval:

Our research was approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

**Reference:**

1. Ronco, C.; Bellomo, R.; Kellum, J.A. Acute kidney injury. *The Lancet* **2019**, *394*, 1949-1964, doi:10.1016/s0140-6736(19)32563-2.
2. Pakula, A.M.; Skinner, R.A. Acute Kidney Injury in the Critically Ill Patient: A Current Review of the Literature. *J Intensive Care Med* **2016**, *31*, 319-324, doi:10.1177/0885066615575699.
3. Bellomo, R.; Kellum, J.A.; Ronco, C. Acute kidney injury. *The Lancet* **2012**, *380*, 756-766, doi:10.1016/s0140-6736(11)61454-2.
4. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* **2012**, *120*, c179-184, doi:10.1159/000339789.
5. Coelho, S.; Cabral, G.; Lopes, J.A.; Jacinto, A. Renal regeneration after acute kidney

injury. *Nephrology (Carlton)* **2018**, *23*, 805-814, doi:10.1111/nep.13256.

6. Wilson, F.P.; Greenberg, J.H. Acute Kidney Injury in Real Time: Prediction, Alerts, and Clinical Decision Support. *Nephron* **2018**, *140*, 116-119, doi:10.1159/000492064.

7. Wu, Y.; Hao, W.; Chen, Y.; Chen, S.; Liu, W.; Yu, F.; Hu, W.; Liang, X. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* **2020**, *42*, 1127-1134, doi:10.1080/0886022X.2020.1843491.

8. Koyner, J.L.; Adhikari, R.; Edelson, D.P.; Churpek, M.M. Development of a Multicenter Ward-Based AKI Prediction Model. *Clin J Am Soc Nephrol* **2016**, *11*, 1935-1943, doi:10.2215/CJN.00280116.

9. Low, S.; Vathsala, A.; Murali, T.M.; Pang, L.; MacLaren, G.; Ng, W.Y.; Haroon, S.; Mukhopadhyay, A.; Lim, S.L.; Tan, B.H., et al. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* **2019**, *20*, 32, doi:10.1186/s12882-019-1206-4.

10. Tomasev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I., et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116-119, doi:10.1038/s41586-019-1390-1.

11. Flechet, M.; Falini, S.; Bonetti, C.; Güiza, F.; Schetz, M.; Van den Berghe, G.; Meyfroidt, G. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Critical Care* **2019**, *23*, 282, doi:10.1186/s13054-019-2563-x.

12. Churpek, M.M.; Carey, K.A.; Edelson, D.P.; Singh, T.; Astor, B.C.; Gilbert, E.R.;

Winslow, C.; Shah, N.; Afshar, M.; Koyner, J.L. Internal and External Validation of a Machine Learning Risk Score for Acute Kidney Injury. *JAMA Netw Open* **2020**, *3*, e2012892, doi:10.1001/jamanetworkopen.2020.12892.

13. Yuan, Q.; Zhang, H.; Deng, T.; Tang, S.; Yuan, X.; Tang, W.; Xie, Y.; Ge, H.; Wang, X.; Zhou, Q., et al. Role of Artificial Intelligence in Kidney Disease. *Int J Med Sci* **2020**, *17*, 970-984, doi:10.7150/ijms.42078.

14. Li, Y.; Xu, J.; Wang, Y.; Zhang, Y.; Jiang, W.; Shen, B.; Ding, X. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* **2020**, *43*, 752-761, doi:10.1002/clc.23377.

15. De Vlieger, G.; Kashani, K.; Meyfroidt, G. Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review. *Curr Opin Crit Care* **2020**, *26*, 563-573, doi:10.1097/MCC.0000000000000775.

16. James, M.T.; Pannu, N.; Hemmelgarn, B.R.; Austin, P.C.; Tan, Z.; McArthur, E.; Manns, B.J.; Tonelli, M.; Wald, R.; Quinn, R.R., et al. Derivation and External Validation of Prediction Models for Advanced Chronic Kidney Disease Following Acute Kidney Injury. *JAMA* **2017**, *318*, 1787-1797, doi:10.1001/jama.2017.16326.

17. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.W.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.A.-O.; Celi, L.A.; Mark, R.A.-O. MIMIC-III, a freely accessible critical care database.

18. Thoral, P.J.; Peppink, J.M.; Driessen, R.H.; Sijbrands, E.J.G.; Kompanje, E.J.O.; Kaplan, L.; Bailey, H.; Kesecioglu, J.; Cecconi, M.; Churpek, M., et al. Sharing ICU

Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. *Crit Care Med* **2021**, 10.1097/CCM.0000000000004916, doi:10.1097/CCM.0000000000004916.

19. Koyner, J.L.; Carey, K.A.; Edelson, D.P.; Churpek, M.M. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit Care Med* **2018**, *46*, 1070-1077, doi:10.1097/CCM.0000000000003123.

**Footnotes**

Contributions: Conceptualization, QL and MH; Methodology, YX; Software, YX; Validation, YZ, and YX.; Formal Analysis, JC and XC; Resources, MH; Data Curation, QL; Writing – Original Draft Preparation, QL; Writing – Review & Editing, MH; Visualization, YX; Supervision, YZ and QL; Project Administration, QL;

**Funding**: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest: Not Applicable.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Acknowledgements: Thank my colleagues in general ICU for cooperating with us in our prospective study. Thanks to two public databases namely MIMIC and AmsterdamUMC for providing important data.

**Supplemental Figure:**

File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.

File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.

File name: **Figure S3. tif**

Title of data: The visual interface for prospective studies of the AKI early diagnosis prediction models.

Description of data: Different patients have independent card display information, in which the last line of the card is the AKI diagnosis. The visual interface was not available to doctors during the prospective study and was used only by the researchers to verify the diagnosis and for sampling verification.

File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

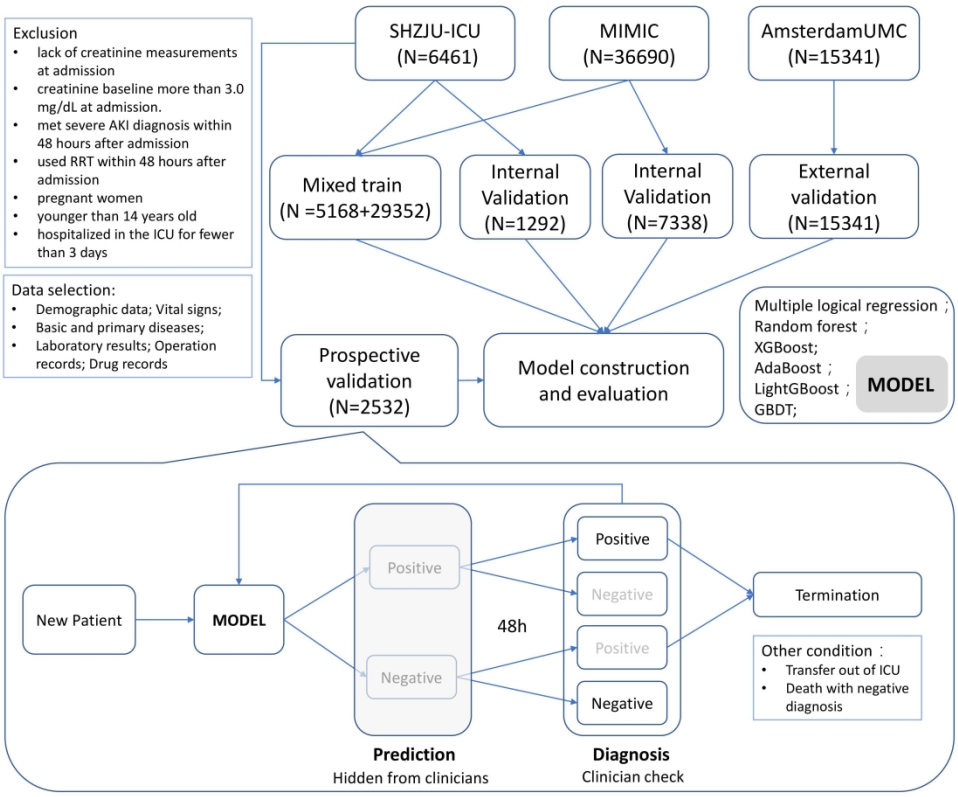
Description of data: Writing norms of this type of research.

**Figure Legend:**

Figure 1: The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.




VISUAL INTERFACE

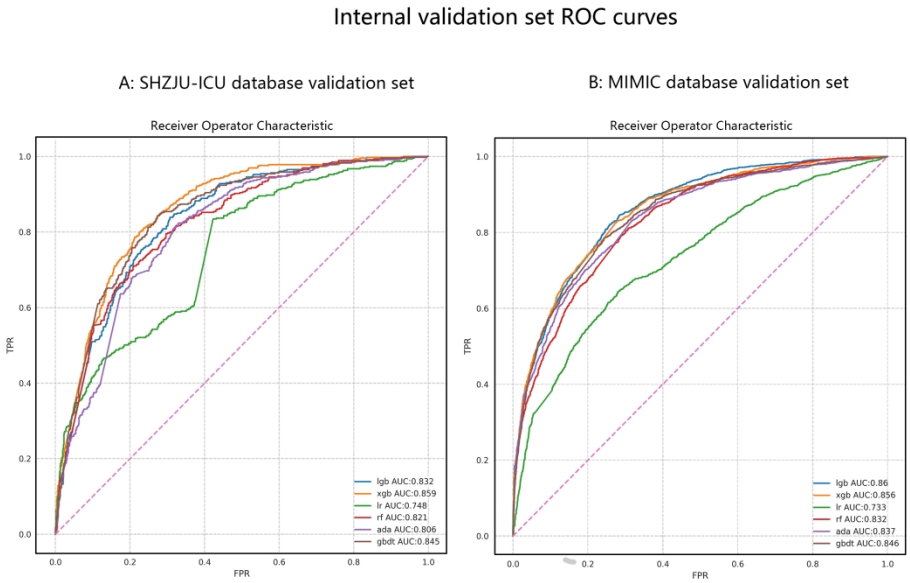



Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

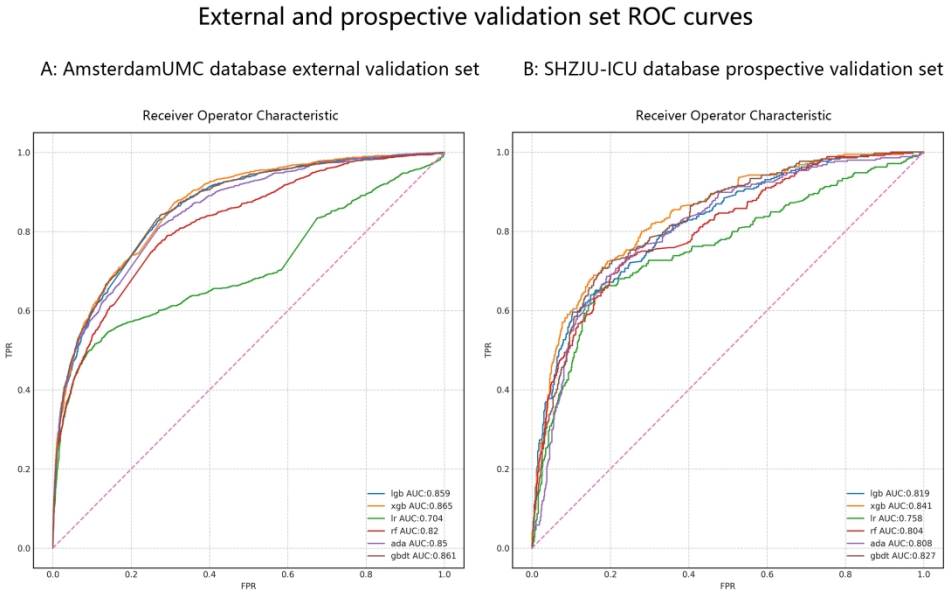
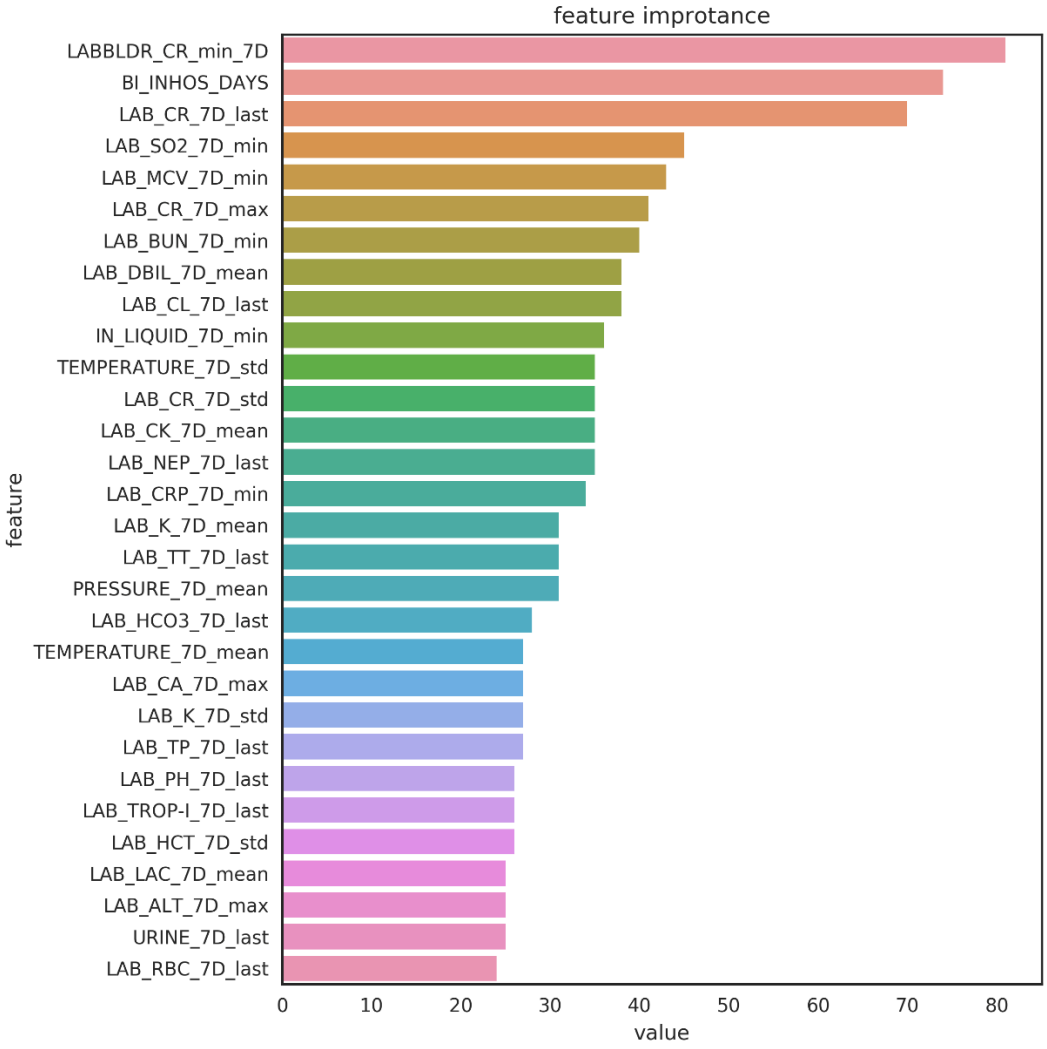



Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.

**Supplemental Figure:**

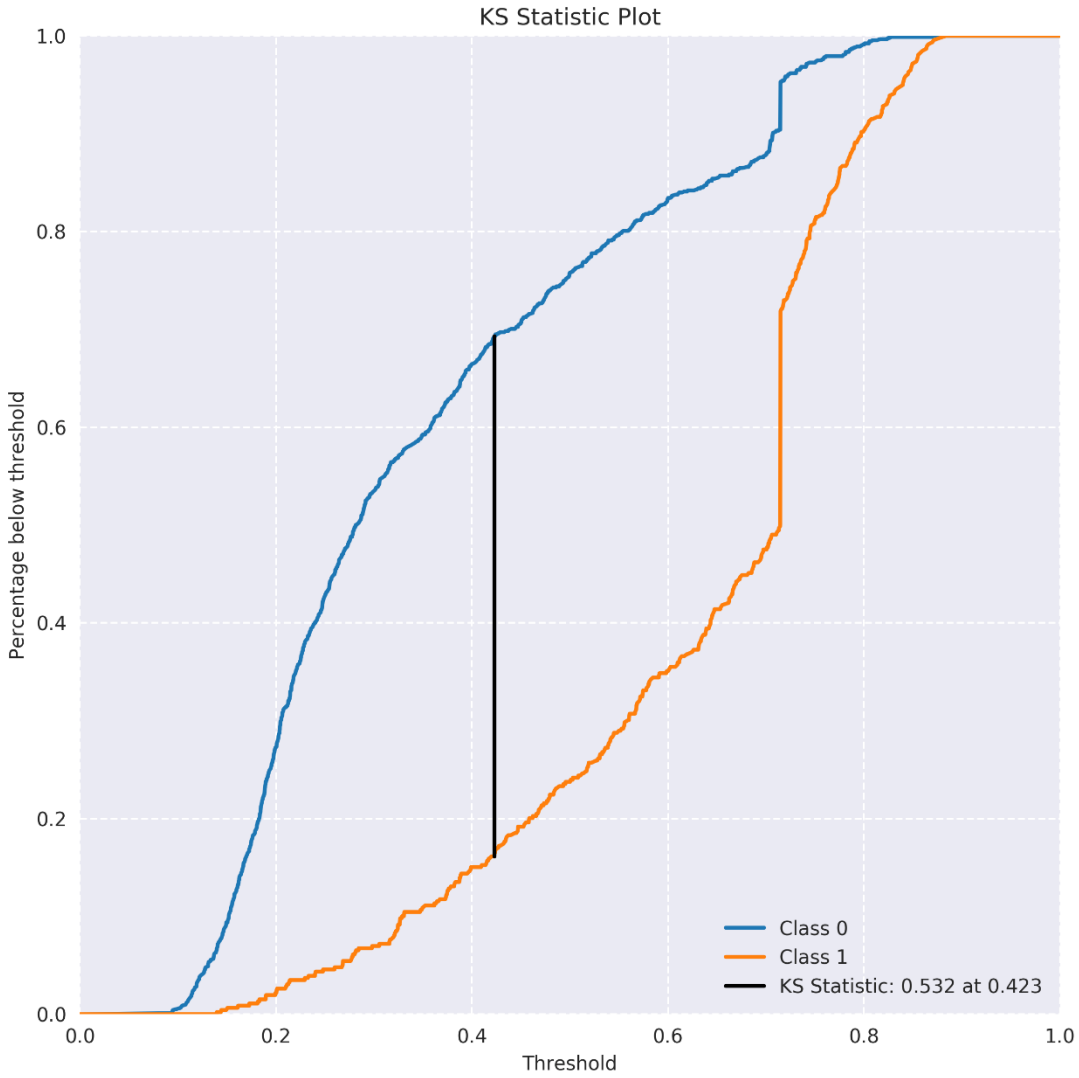
File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.




File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.




File name: **Figure S3. tif**

Title of data: The visual interface for prospective studies of the AKI early diagnosis prediction models.

Description of data: Different patients have independent card display information, in which the last line of the card is the AKI diagnosis. The visual interface was not available to doctors during the prospective study and was used only by the researchers to verify the diagnosis and for sampling verification.


File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

| | SHZJU-ICU | MIMIC | AmsterdamUMC | Including in Model |
|---|---|---|---|---|
| **Age** | √ | √ | √ | √ |
| **Gender** | √ | √ | √ | √ |
| **Race** | √ | √ | × | × |
| **Primary disease** | √ | √ | × | × |
| **Comorbidity** | √ | √ | × | × |
| **Ventilation** | √ | √ | √ | √ |
| **Operation** | √ | √ | × | × |
| **ICU hours** | √ | √ | √ | √ |
| **Survived** | √ | √ | √ | √ |
| **Vital sign** | | | | |
| **Temperature** | √ | √ | √ | √ |
| **Systolic pressure** | √ | √ | √ | √ |
| **Diastolic pressure** | √ | √ | √ | √ |
| **Respiratory rate** | √ | √ | √ | √ |
| **Heart rate** | √ | √ | √ | √ |
| **Oxygen saturation** | √ | √ | √ | √ |
| **Urine** | √ | √ | √ | √ |
| **GCS score** | √ | × | × | × |
| **Laboratory results** | | | | |

| | | | | |
|---|---|---|---|---|
| **Scr** | √ | √ | √ | √ |
| **BUN** | √ | √ | √ | √ |
| **RBC** | √ | √ | √ | √ |
| **MCV** | √ | √ | √ | √ |
| **Hb** | √ | √ | √ | √ |
| **WBC** | √ | √ | √ | √ |
| **NEUT** | √ | √ | √ | √ |
| **PLT** | √ | √ | √ | √ |
| **HCV** | √ | √ | × | × |
| **TBLB** | √ | √ | √ | √ |
| **DBLB** | √ | √ | √ | √ |
| **IBLB** | √ | √ | √ | √ |
| **CRP** | √ | √ | √ | √ |
| **PCT** | √ | × | × | × |
| **Serum kalium** | √ | √ | √ | √ |
| **Serum natrium** | √ | √ | √ | √ |
| **Serum chlorine** | √ | √ | √ | √ |
| **PT** | √ | √ | √ | √ |
| **APTT** | √ | √ | √ | √ |
| **INR** | √ | √ | √ | √ |
| **CK** | √ | √ | √ | √ |
| **LDH** | √ | √ | √ | √ |
| **Troponin** | √ | √ | √ | √ |
| **Blood glucose** | √ | √ | √ | √ |
| **PH** | √ | √ | √ | √ |
| **Lactic acid** | √ | √ | √ | √ |
| **Anion gap** | √ | √ | √ | √ |
| **ABC** | √ | √ | √ | √ |
| **SBC** | √ | √ | √ | √ |
| **$PaCO_2$** | √ | √ | √ | √ |
| **$PaO_2$** | √ | √ | √ | √ |

SHZJU-ICU: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; MIMIC: Medical Information Mart for Intensive Care; APTT: Activated partial thromboplastin time; INR: International normalized ratio; LDH: Lactate dehydrogenase; $PaCO_2$: Partial pressure of carbon dioxide; $PaO_2$ Partial pressure of oxygen; MCV: Mean corpuscular volume; PH: Potential of hydrogen; WBC: White blood cell count; NEUT: Proportion of neutrophils; CK: Creatine kinase; Scr: Serum Creatinine; BUN: Blood urea nitrogen; RBC: Red blood cell count; Hb: Hemoglobin; PT: Prothrombin time; TBLB: Total bilirubin; DBLB: Direct bilirubin; IBLB: Indirect bilirubin; CRP: C-reactive protein; HCV: Hematocrit value; PCT: Procalcitonin; ABC: Actual Bicarbonate Radical; SBC: Standard Bicarbonate Radical; PLT: Platelet count

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | SHZJU-ICU Prospective Negative (N=2188) | SHZJU-ICU Prospective Positive (N=344) | P-value |
|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 61.3 [47.5-73.4] | 62.1 [51.2-70.5] | 0.03 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 1365 (62.4) | 207 (60.2) | 0.438 |
| **Race** | | | | | | |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 2 (0.001) | 0 | 1 |
| **Black** | 1 (0.02) | 0 | 1 | 0 | 0 | 1 |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 2185 (99.9) | 344(100) | 1 |
| **Other** | 0 | 1 (0.1) | 1 | 1 (0.001) | 0 | 1 |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 0.8 (0.32) | 0.9 (0.35) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (11) | 20 (10) | 0.007 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 230 (10.5) | 26 (7.6) | 0.102 |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 15 (0.7) | 3 (0.8) | 0.726 |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 98 (4.5) | 19 (5.7) | 0.403 |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 142 (6.5) | 29 (8.6) | 0.203 |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 324(14.8) | 56 (16.4) | 0.466 |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 1102 (50.4) | 239 (69.5) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 1234(56.4) | 160 (46.7) | 0.001 |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 58.7 [34.2-160.4] | 145.6 [68.5-314.8] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 1868 (85.4) | 178 (51.9) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

| 1Section/Topic | | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L80 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L84-86 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P5, L104 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L116 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L119 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L126 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P6, L159 |
| | 10e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | NA |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |

| | | | | |
|---|---|---|---|---|
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 3a | );V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 3b | );V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 |
| | 3c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, Table2 |
| Model development | 4a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 4b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes |
| Model specification | 5a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 5b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | );V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | );V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L236 |
| Interpretation | 9a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L217 |
| | 9b | );V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L198 |
| Implications | 20 | );V | Discuss the potential clinical use of the model and implications for future research. | Yes, P9, L193 |
| **Other information** | | | | |
| Supplementary information | 21 | );V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | );V | Give the source of funding and the role of the funders for the present study. | NA |

BMJ Open

# Prospective Validation for Severe Acute Kidney Injury Predicting Model based on transcontinental databases



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2021-054092.R1 |
| Article Type: | Original research |
| Date Submitted by the Author: | 23-Sep-2021 |
| Complete List of Authors: | Liang, Qiqiang; Zhejiang University School of Medicine Second Affiliated Hospital<br>Xu, Yongfeng; HealSci Technology Co<br>Zhou, Yu; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Xinyi; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Juan; Zhejiang University School of Medicine Second Affiliated Hospital<br>Huang, Man; Zhejiang University School of Medicine Second Affiliated Hospital, |
| Primary Subject Heading: | Intensive care |
| Secondary Subject Heading: | Diagnostics |
| Keywords: | Acute renal failure < NEPHROLOGY, INTENSIVE & CRITICAL CARE, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Prospective Validation for Severe Acute Kidney Injury Predicting Model based on transcontinental databases

Qiqiang Liang[1]; Yongfeng Xu[2]; Yu Zhou[1]; Xinyi Chen[1], MM; Juan Chen[1]; Man Huang[1*]

[1]General Intensive Care Unit, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

[2]BSc, HealSci Technology Co., Ltd, Beijing China.

*Corresponding author: Man Huang, MD PhD.

Department: General Intensive Care Unit

Hospital: Second Affiliated Hospital of Zhejiang University School of Medicine

Address: No. 1511, Jianghong Road, Bingjiang District, Hangzhou City, Zhejiang Province

Tel: + 86 571 89713427

Fax: + 86 571 89713427

E-mail: huangman@zju.edu.cn.

ORCID: 0000-0002-0842-2254

**Abstract:**

Objectives: There are many studies of acute kidney injury (AKI) diagnosis models lack of external validation and prospective validation. We constructed the models using three databases to predict severe AKI within 48 hours in intensive care unit (ICU) patients.

Design: A retrospective and prospective cohort study.

Setting: We studied critically ill patients in our database (SHZJU-ICU) and two other public databases, the Medical Information Mart for Intensive Care (MIMIC) and AmsterdamUMC databases, including basic demographics, vital signs, and laboratory results. We predicted the diagnosis of severe AKI in patients in the next 48 hours using machine-learning algorithms with the three databases. Then, we carried out real-time severe AKI prediction in the prospective validation study at our centre for one year.

Participants: All patients included in three databases with uniform exclusion criteria.

Primary and secondary outcome measures: Effect evaluation index of prediction models.

Results: We included 58492 patients, and a total of 5257 (9.0%) patients met the definition of severe AKI. In the internal validation of the SHZJU-ICU and MIMIC databases, the best area under the receiver operating characteristic curve (AUROC) of the model was 0.86. The external validation results by AmsterdamUMC database were also satisfactory, with the best AUROC of 0.86. A total of 2532 patients were admitted to the center for prospective validation; 358 positive results were predicted, and 344 patients were diagnosed with severe AKI, with the best sensitivity of 0.72, the specificity of 0.80, and the AUROC of 0.84.

Conclusion: The prediction model of severe AKI exhibits promises as a clinical application based on dynamic vital signs and laboratory results of multi-center databases with prospective and external validation.

**Keywords**: Machine learning; Acute kidney injury; Real-time prospective validation; External validation.

**Strengths and limitations of this study**:

An important verification step from artificial intelligence research to clinical use;

Three large database containing different national populations and regions;

The prediction model with excellent performance in complete data validation;

Differences in the samples proportion of three databases;

The dimensions of variables are not rich enough.

## Introduction:

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalization and mortality[1]. AKI is divided into three types according to the various aetiologies: prerenal (renal hypoperfusion), intrarenal (vascular, glomerular, or tubulointerstitial lesions), and postrenal (urinary tract obstruction)[2]. Although nearly all diseases associated with ICU admission may cause AKI, acute tubular necrosis (ATN) and prerenal azotaemia are the most common causes[3].

All AKI diagnostic criteria including the latest Kidney Disease: Improving Global Outcomes (KDIGO) standard are currently based on the creatinine level and urine volume[4]. However, the increase in the creatinine level or decrease in the urine volume lags the onset of AKI[2]. Many studies have suggested that early diagnosis and treatment of reversible AKI can reduce mortality[5]. Therefore, the creatinine level and urine volume are not satisfactory to meet clinical diagnostic demands. Consequently, many researchers have tried to develop an early warning model by analysing the risk factors for AKI[6].

Patient complications, such as diabetes, hypertension, cardiovascular disease, chronic liver disease, sepsis, and trauma, are identified as important risk factors for AKI[7]. The AKI prediction model and scoring system developed based on high-risk factors has gradually become the focus of research considering the lower clinical application threshold compared with that of new biomarkers[6]. Although most previous prediction models use the multiple logistic regression model, a variety of AKI prediction models based on machine learning have resulted in satisfactory outcomes[6]. Since the first AKI prediction model study based on artificial intelligence was published in 2016, researchers have built more than 20 published AKI prediction models successively by using local or multi-centre databases[6,8-14]. The results indicate that these models can predict the occurrence of AKI and the need for renal replacement therapy (RRT) within 24 or 48 hours, with accuracies ranging from 81% to 97%[6,15]. In addition, many studies have focused on subspecialized conditions, including cardiac surgery, trauma, and burns[14-16]. However, the common defect in these studies is the lack of external validation and prospective validation, which causes the prediction model to deviate from the clinical scenarios and limits extrapolation beyond the scope of the data.

In this study, we built models to predict AKI within 48 hours in critically ill patients by using transcontinental three databases. Then, we evaluated the clinical effect of the model through a one-year prospective validation at our centre.

## Methods :

### Study design and setting

We collected patients using three ICU databases and prospectively validated the models in our centre. The first database was our centre general ICU database (SHZJU-ICU) of the Second Affiliated Hospital

of Zhejiang University School of Medicine, an academic teaching hospital. Since its establishment in 2017, it has included 12000 ICU patients' data and is updated daily. The Medical Information Mart for Intensive Care (MIMIC) III database, the second one, is an open ICU database provided by the Massachusetts Institute of Technology and includes nearly 60000 ICU patients from North America[17]. Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23000 patients admitted to ICUs in parts of Europe[18]. The research flow chart is shown in Figure 1.

**Study definition**

In this study, the diagnosis of AKI was confirmed based on three stages according to the KDIGO criteria[4]. We defined the patients who met the KDIGO AKI II and III criteria as severe AKI groups and the others as negative groups. We excluded patients with lack of creatinine measurements during admission, patients with creatinine baseline more than 3.0 mg/dL at admission, patients who met severe AKI diagnosis within 24 hours, and patients who used RRT within 48 hours after admission[19]. In addition, we excluded pregnant women, patients younger than 14 years old, and patients hospitalized in the ICU for fewer than 48 hours. After the patient was admitted to ICU, we performed a prediction every 24 hours and recorded a prediction time. If the patient was diagnosed with severe AKI within 48 hours, the predictive time was defined as a positive predictive point, and the others were defined as a negative point.

**Data Collection**

The variables included demographic data, vital signs, basic and primary diseases, laboratory results, important operation records, and drug records. Comorbidity included hypertension, diabetes, cardiopathy, liver disease, and malignant tumors. The primary disease was the main cause of admission to the ICU

following the ICD-10 codes. The vital signs and clinical laboratory results were transformed into different variables according to the average, variance, maximum, minimum, and final value before diagnosis. We use a method similar to the forward incremental method in the multivariate logic regression model, that is, the combination of embedded feature selection and forward addition for feature selection. First of all, all variables are trained in the model, then list by variables importance. variables are added to the model one by one according to the variable importance. a variable is retained if it causes the AUC growth to be greater than 0.01, otherwise delete it. We transformed the MIMIC and AmsterdamUMC databases according to our centre database structure, unifying the unit and diagnostic codes. We deleted variables missing more than 50%. Variables missing more than 30% but less than 50% are listed to clinicians who determine the potential correlation between these variables and AKI. We carry out multiple interpolation for these variables which clinicians require to be retained, and the others deleted. Variables missing less than 30% are fill in multiple interpolation. All Missing data between three databases and values included in the model shown in supplementation file Table S1.

**Model construction and external validation**

The ratio of the training and internal validation sets was 4:1. The SHZJU and MIMIC databases training sets were mixed into a new training set. There were more negative data than positive data, so we randomly sampled the negative datasets and constructed a new data subset with a sampling ratio of positive and negative data of 1:5. We used multiple logistic regression, random forest, XGBoost, AdaBoost, LightGBoost, gradient boosting decision tree (GBDT), and debug to assess the variables and model-related parameters by the fivefold cross-validation method. After the models were built, we used the SHZJU-ICU and MIMIC test sets for internal validation of the model and the AmsterdamUMC database for external validation. The most appropriate cut-off value was determined according to the K-S

curve. The prediction model represents the results of each prediction with a probability between 0 to 1.0. We define results more than 0.4 as high-risk, that is, positive results, and the rest as negative results. Through internal validation and external validation, we calibrated the model by adjusting the super-parameters and using the Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. The model was visualized and analyzed by a visual programming scheme. All model building and validation processes were performed in Python 3.6.

**Prospective validation**

The prospective research period was 2020.01.01 to 2020.12.31. We collected real-time data when patients were admitted to the ICU, transformed the data according to the requirements, and formed a complete sample for the prediction model after passing the integrity test. We had established a visualization scheme and allowed researchers to review the predictions daily. The daily prediction results were not publicly accessible during the study to avoid affecting clinicians' decisions, but the diagnosis results were available to the researchers as visual graphics. We sampled the 20% predicted data every month and deleted samples with more than 50% missing values to ensure data correctness. When a patient has the following conditions, AKI prediction system will end the patient's prospective prediction: A. a positive diagnosis; B. Transfer out of ICU or death with negative diagnosis. All diagnosis of severe AKI needs to be reviewed by two ICU attending physicians independently, and if the they have different opinion, the third one will be appealed.

**Patient and Public Involvement:**

The information of cases in three databases was in a state of complete desensitization in the process of building the model. During the prospective study, all the patients signed an informed consent form at the beginning of admission to ICU. The real-time data discussed and used by the study members only,

and were not made public during the study period. The study was evaluated and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine as study number 2019-078. All data were anonymized before the authors accessed them for the purpose of this study. According to the inclusion criteria and exclusion criteria, we selected 58492 patients from three databases who met the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36690 patients from the MIMIC database, and 15341 patients from the AmsterdamUMC database.

**Statistical analysis:**

The population characteristics were reported as the medians and inter-quartile ranges (IQRs) for skewed data and the means and standard deviations for normally distributed data. The independent sample T-test was used for normally distributed data, and the rank-sum test was used for the rest. Dichotomous variables were assessed by the chi-square test, and a P value less than 0.05 was considered statistically significant. The non-normally distributed data were analyzed by exponential transformation and logarithmic transformation. The effect of the model was evaluated by parameters such as the AUROC, accuracy, specificity, and F1-score.

**Results:**

A total of 5257 (9.0%) patients met the definition of severe AKI (11.8% in SHZJU-ICU, 7.6% in MIMIC, and 10.9% in AmsterdamUMC). The distributions of age and sex in the three centers were similar, but the differences in race were large. Asian patients accounted for more than 99% in the SHZJU-ICU database, and only approximately 2.5% in the MIMIC database. White people accounted for more than 70% of the MIMIC database. In addition, patients from the MIMIC database had a higher incidence of the tumor, liver cirrhosis, diabetes, and hypertension. Patients in the AmsterdamUMC and SHZJU-

ICU databases had a higher proportion of mechanical ventilation and overall survival rate. Severe AKI patients had longer ICU hospital stays and higher mortality. More details are presented in Table 1.

Table 1: Clinical Demographics and Outcomes in patients with or without severe acute kidney injury.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | MIMIC Negative (N=33879) | MIMIC Severe AKI (N=2811) | P-value | AmsterdamUMC Negative (N=13661) | AmsterdamUMC severe AKI (N=1680) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 62.0 [45.0-76.0] | 64.0 [52.0-76.0] | <0.001 | 64.5 [54.5,74.5] | 64.3 [52.5,78.5] | 0.04 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 18976 (56.0) | 1524 (54.2) | 0.06 | 8920 (65.3) | 1054 (63.9) | 0.03 |
| **Race** | | | | | | | / | / | / |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 23796 (70.2) | 2005 (71.3) | 0.23 | / | / | / |
| **Black** | 1 (0.02) | 0 | 1 | 2455 (7.2) | 199 (7.1) | 0.45 | / | / | / |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 918 (2.7) | 70 (2.5) | 0.54 | / | / | / |
| **Other** | 0 | 1 (0.1) | 1 | 6710 (19.8) | 537 (19.1) | 0.34 | / | / | / |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 1.1 (0.41) | 1.1 (0.50) | 0.04 | 1.2 (0.35) | 1.1 (0.38) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (12) | 23 (12) | <0.001 | 18 (13) | 23 (10) | <0.001 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 6652 (19.6) | 752 (26.8) | <0.001 | / | / | / |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 1180 (3.5) | 426 (15.2) | <0.001 | / | / | / |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 4840 (14.3) | 336 (12.0) | 0.001 | / | / | / |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 7250 (21.4) | 789 (28.1) | <0.001 | / | / | / |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 16328 (48.2) | 1513 (53.8) | <0.001 | / | / | / |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 13009 (38.4) | 1503(53.5) | <0.001 | 7718(56.5) | 1248(74.3) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 12455 (36.7) | 1430 (50.8) | <0.001 | / | / | / |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 51.0 [28.5-110.3] | 103.4 [50.1-261.3] | <0.001 | 24.0 [19.8,63.5] | 142.0 [45.4,394.2] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 25648 (75.7) | 1356 (48.2) | <0.001 | 10942(80.1) | 1033 (61.5) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile

Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

There were significant differences in the important parameters of the variables among the different models (see Figure S1). However, the trend of the creatinine level in the past week was still an important variable, followed by urine volume, blood urea nitrogen level, temperature, and length of ICU stay. The cut-off value used to distinguish between a negative and positive prediction was determined by the K-S curve, with value of 0.423 (see Figure S2). The GBDT model had the best prediction effect in the test set, followed by XGBoost and LightGBoost. In the two central internal validation sets, the two best-performing machine learning algorithms with great AUROC are LightGBoost (SHZJU-ICU of 83.2%, MIMIC of 86.0%) and XGBoost (SHZJU-ICU 85.9%, MIMIC 85.6%), as detailed in Figure 2. Overall, the sensitivity (SHZJU-ICU 0.84, MIMIC 0.83) and the negative predictive value (SHZJU-ICU 0.90, MIMIC 0.90) of the predictive model were high, but the specificity was general (SHZJU-ICU 0.79, MIMIC 0.75), as shown in Table 2. In the external validation based on AmsterdamUMC database, the overall model validation effect results were satisfactory, and XGBoost had the best performance, with an AUROC of 0.84, as shown in Figure 2 and Table 2. We built a visual prediction interface based on the prediction model (supplementation file Figure S3).

Table 2: Model validation results by three databases with machine learning algorithm

| *Model* | *AUROC* | *Accuracy* | *Sensitivity* | *~~Sepcificity~~ Specificity* | *PPV* | *NPV* | *F1* |
|---|---|---|---|---|---|---|---|
| ***Internal validation with SHZJU-ICU database*** | | | | | | | |
| ***Logistic regression*** | *0.748* | *0.662* | *0.834* | *0.576* | *0.496* | *0.874* | *0.622* |
| ***LightGBoost*** | *0.832* | *0.741* | *0.839* | *0.692* | *0.576* | *0.896* | *0.683* |
| ***GBDT*** | *0.845* | *0.765* | *0.843* | *0.725* | *0.606* | *0.902* | *0.705* |
| ***AdaBoost*** | *0.806* | *0.721* | *0.824* | *0.67* | *0.555* | *0.884* | *0.663* |
| ***Random Forest*** | *0.821* | *0.763* | *0.71* | *0.789* | *0.627* | *0.845* | *0.666* |
| ***XGBoost*** | ***0.859*** | *0.779* | *0.81* | *0.763* | *0.631* | *0.889* | *0.709* |
| ***Internal validation with MIMIC database*** | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ***Logistic regression*** | *0.733* | *0.695* | *0.643* | *0.72* | *0.535* | *0.801* | *0.584* |
| ***LightGBoost*** | ***0.86*** | *0.768* | *0.822* | *0.741* | *0.613* | *0.893* | *0.702* |
| ***GBDT*** | *0.846* | *0.765* | *0.786* | *0.755* | *0.616* | *0.876* | *0.691* |
| ***AdaBoost*** | *0.837* | *0.732* | *0.831* | *0.683* | *0.567* | *0.89* | *0.674* |
| ***Random Forest*** | *0.832* | *0.738* | *0.791* | *0.712* | *0.578* | *0.872* | *0.668* |
| ***XGBoost*** | *0.856* | *0.758* | *0.833* | *0.721* | *0.598* | *0.895* | *0.695* |
| ***External validation with AmsterdamUMC database*** | | | | | | | |
| ***Logistic regression*** | *0.704* | *0.767* | *0.516* | *0.893* | *0.706* | *0.787* | *0.596* |
| ***LightGBoost*** | *0.859* | *0.763* | *0.827* | *0.731* | *0.606* | *0.894* | *0.7* |
| ***GBDT*** | *0.861* | *0.764* | *0.84* | *0.727* | *0.606* | *0.901* | *0.704* |
| ***AdaBoost*** | *0.85* | *0.755* | *0.813* | *0.726* | *0.597* | *0.886* | *0.689* |
| ***Random Forest*** | *0.82* | *0.743* | *0.77* | *0.729* | *0.587* | *0.864* | *0.666* |
| ***XGBoost*** | ***0.865*** | *0.75* | *0.873* | *0.688* | *0.584* | *0.916* | *0.7* |
| ***Prospective validation with SHZJU-ICU*** | | | | | | | |
| ***Logistic regression*** | *0.758* | *0.772* | *0.648* | *0.834* | *0.662* | *0.826* | *0.655* |
| ***LightGBoost*** | *0.819* | *0.796* | *0.596* | *0.895* | *0.74* | *0.816* | *0.66* |
| ***GBDT*** | *0.827* | *0.781* | *0.706* | *0.818* | *0.66* | *0.848* | *0.683* |
| ***AdaBoost*** | *0.808* | *0.766* | *0.686* | *0.805* | *0.638* | *0.837* | *0.661* |
| ***Random Forest*** | *0.804* | *0.755* | *0.715* | *0.775* | *0.613* | *0.845* | *0.66* |
| ***XGBoost*** | ***0.841*** | *0.779* | *0.724* | *0.807* | *0.652* | *0.854* | *0.686* |

*AUROC*: Area Under Receiver Operating Characteristic; *PPV*: Positive Predictive Value; *NPV*: Negative Predictive Value; *SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *GBDT*: Gradient Boosted Decision Tree.

According to the inclusion and exclusion criteria, we delete 267 patients among 94 patients with creatinine baseline more than 3.0 mg/dL at admission, 39 patients met severe AKI diagnosis within 24 hours, and 26 patients who used RRT within 48 hours after admission, 108 patients hospitalized in the ICU for fewer than 48 hours. A total of 2532 patients were admitted to our centre for prospective validation, and the prediction model made 16858 times predictions. In the prospective cohort, there was no significant difference in age, gender, baseline creatinine and urea nitrogen, and complications. The proportion of mechanical ventilation and the ICU stay time in AKI patients were longer with higher mortality. Above all, there was no significant difference between the prospective and the retrospective cohort. More detail sees in supplementation file Table S2. In the end, 358 positive results were predicted,

and the rest were negative results. There are 344 patients with severe AKI were diagnosed and the prediction accuracy was 83.5%. The model with the highest area under the curve was XGBoost, 0.84 with the best sensitivity of 0.72, the specificity of 0.80. The results of the prospective study are similar to those of the external validation of the model, and are relatively stable. More detail is presented in Figure 3 and Table 2.

**Discussion:**

In this study, we built predictive models by machine learning to predict the incidence of severe AKI with three databases in different regions and in the next 48 hours. After internal and external validation, prospective validation over one year was carried out to verify the model effects. The three databases come from three countries that are in Asia, Europe, and North America, which proves that the model is universal to some extent.

Despite the huge amount of data, many databases are still not suitable for prospective research because they are not updated promptly. Tomasev, N and colleagues have provided research on AKI prediction models with a large amount of data[10]. The study covered 703782 adult patients with 6 billion individual items, including 620000 elements. In this study, a depth neural network model was used for real-time prediction. A total of 55.8% of severe AKI patients were predicted within the first 48 hours, although each accurate prediction was accompanied by two mispredictions[10]. This study provided a new scheme for real-time prediction and indicated that we should prospectively evaluate and independently validate models to explore their effectiveness. In a prospective study, Flechet, M et al compared an AKI prediction model with clinicians in 252 patients and found that the clinical effect of the random forest model for predicting AKI-II/III was equivalent to that of clinicians. Our prediction

model graphical visualization of the model was installed in the centre's database for better usage. In addition, our database is updated daily to achieve daily predictions and present the results to researchers. In the prospective validation of our study, the stability of the prediction model confirmed its promise, which provides a basis for future research.

There are many studies of artificial intelligence for predicting the occurrence of AKI, but most of them are single-centre studies, and the extrapolation effect has been controversial. In 2016, Koyner, J, and his colleague published the first study of an AKI prediction model based on multi-centre data. In all 202961 patients, 17541 (8.6%) had AKI, 4251 (3.5%) had AKI-II, and 1242 (0.6%) had AKI-III. A multivariate logistic regression model was used to predict AKI in this study with an AUROC of 0.74 (95% CI=0.74). With the classification of AKI, the AUROC of the prediction model gradually increased to 0.84[8]. Subsequently, the study team used a new machine learning algorithm, to build a more accurate model to predict the occurrence of AKI-II, with an AUROC of 0.9 within 24 hours and 0.87 within 48 hours[19]. Recently, the research team included data from two other centres, namely, LUMC (N=200613) and NUS (N=246895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85 to 0.86, suggesting that the artificial intelligence model has stable predictive ability[12]. This series of studies included many data points, suggesting the feasibility of artificial intelligence in the diagnosis of AKI, but the proportion of positive patients (3.5%) and ICU patients (30%) was too low to properly predict AKI. Our research is similar to the above. The SAHZJU-ICU database is a single-centre database representing south-eastern China, and the MIMIC database is a well-known open ICU database in the United States. The AmsterdamUMC database is a public database located in Europe. The population structure and diseases in the three databases are complete but different in the distribution of complications and race. Therefore, it provides a prediction model with unparalleled stability compared with other

studies.

Limitations:

This retrospective multi-centre study was unable to carry out more clinical feature mining and comparison because of different data structures. The differences between the three databases partly reflect some demographic differences between Europe, the United States and China, resulting in a decline in the accuracy of the prediction model. There are some differences in the number of patients included in the three databases, which may affect the choice of variables. As a result of the study design, we deleted patients with ICU hospitalization of less than 48 hours, which may result in the exclusion of most relatively mild patients and may reduce false positives. Second, in the prospective data study in 2020, there may be deviations in the inclusion of patients in the centre, thus affecting the interpretation of the follow-up prospective results. Finally, given the low incidence of severe AKI and the great difference in the proportion of positive and negative samples, the data may be accidental. The diagnostic performance of severe AKI is good with the sensitivity as high as 0.85 in model construction and external validation. however, the sensitivity decreases to 0.72 in the prospective validation, and the overall PPV effect is general. Our model seems to be superior to diagnostic non-AKI patients rather than AKI because of the proportion of positive data that we include. A large number of negative data will increase the specificity and reduce the sensitivity. In the retrospective study, we reduced the proportion of negative data by randomization but retain all date in prospective phase with the sensitivity decreases. We believe that such results are still acceptable and need to be viewed by the reader as a whole.

**Conclusion:**

Based on databases of patients of different races from different countries, we constructed stable machine learning models to predict the occurrence of AKI in the next 48 hours. Prospective validation through the implementation of an updated local database is an effective exploration of further research.

Data Availability Statement:

The model code can be obtained by email if readers need it, but we cannot guarantee that all the code will be provided. Two public databases can be applied from the official website of their respective databases.

Ethics approval:

Our research was approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

Competing interests for authors:

The authors declare that they have no competing interests.

**Reference:**

1. Ronco, C.; Bellomo, R.; Kellum, J.A. Acute kidney injury. *The Lancet* **2019**, *394*, 1949-1964, doi:10.1016/s0140-6736(19)32563-2.
2. Pakula, A.M.; Skinner, R.A. Acute Kidney Injury in the Critically Ill Patient: A Current Review of the Literature. *J Intensive Care Med* **2016**, *31*, 319-324, doi:10.1177/0885066615575699.
3. Bellomo, R.; Kellum, J.A.; Ronco, C. Acute kidney injury. *The Lancet* **2012**, *380*, 756-766, doi:10.1016/s0140-6736(11)61454-2.

4. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* **2012**, *120*, c179-184, doi:10.1159/000339789.

5. Coelho, S.; Cabral, G.; Lopes, J.A.; Jacinto, A. Renal regeneration after acute kidney injury. *Nephrology (Carlton)* **2018**, *23*, 805-814, doi:10.1111/nep.13256.

6. Wilson, F.P.; Greenberg, J.H. Acute Kidney Injury in Real Time: Prediction, Alerts, and Clinical Decision Support. *Nephron* **2018**, *140*, 116-119, doi:10.1159/000492064.

7. Wu, Y.; Hao, W.; Chen, Y.; Chen, S.; Liu, W.; Yu, F.; Hu, W.; Liang, X. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* **2020**, *42*, 1127-1134, doi:10.1080/0886022X.2020.1843491.

8. Koyner, J.L.; Adhikari, R.; Edelson, D.P.; Churpek, M.M. Development of a Multicenter Ward-Based AKI Prediction Model. *Clin J Am Soc Nephrol* **2016**, *11*, 1935-1943, doi:10.2215/CJN.00280116.

9. Low, S.; Vathsala, A.; Murali, T.M.; Pang, L.; MacLaren, G.; Ng, W.Y.; Haroon, S.; Mukhopadhyay, A.; Lim, S.L.; Tan, B.H., et al. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* **2019**, *20*, 32, doi:10.1186/s12882-019-1206-4.

10. Tomasev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I., et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116-119, doi:10.1038/s41586-019-1390-1.

11. Flechet, M.; Falini, S.; Bonetti, C.; Güiza, F.; Schetz, M.; Van den Berghe, G.; Meyfroidt, G. Machine learning versus physicians' prediction of acute kidney injury in critically ill

adults: a prospective evaluation of the AKIpredictor. *Critical Care* **2019**, *23*, 282, doi:10.1186/s13054-019-2563-x.

12. Churpek, M.M.; Carey, K.A.; Edelson, D.P.; Singh, T.; Astor, B.C.; Gilbert, E.R.; Winslow, C.; Shah, N.; Afshar, M.; Koyner, J.L. Internal and External Validation of a Machine Learning Risk Score for Acute Kidney Injury. *JAMA Netw Open* **2020**, *3*, e2012892, doi:10.1001/jamanetworkopen.2020.12892.

13. Yuan, Q.; Zhang, H.; Deng, T.; Tang, S.; Yuan, X.; Tang, W.; Xie, Y.; Ge, H.; Wang, X.; Zhou, Q., et al. Role of Artificial Intelligence in Kidney Disease. *Int J Med Sci* **2020**, *17*, 970-984, doi:10.7150/ijms.42078.

14. Li, Y.; Xu, J.; Wang, Y.; Zhang, Y.; Jiang, W.; Shen, B.; Ding, X. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* **2020**, *43*, 752-761, doi:10.1002/clc.23377.

15. De Vlieger, G.; Kashani, K.; Meyfroidt, G. Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review. *Curr Opin Crit Care* **2020**, *26*, 563-573, doi:10.1097/MCC.0000000000000775.

16. James, M.T.; Pannu, N.; Hemmelgarn, B.R.; Austin, P.C.; Tan, Z.; McArthur, E.; Manns, B.J.; Tonelli, M.; Wald, R.; Quinn, R.R., et al. Derivation and External Validation of Prediction Models for Advanced Chronic Kidney Disease Following Acute Kidney Injury. *JAMA* **2017**, *318*, 1787-1797, doi:10.1001/jama.2017.16326.

17. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.W.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.A.-O.; Celi, L.A.; Mark, R.A.-O. MIMIC-III, a freely accessible critical

care database.

18. Thoral, P.J.; Peppink, J.M.; Driessen, R.H.; Sijbrands, E.J.G.; Kompanje, E.J.O.; Kaplan, L.; Bailey, H.; Kesecioglu, J.; Cecconi, M.; Churpek, M., et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. *Crit Care Med* **2021**, 10.1097/CCM.0000000000004916, doi:10.1097/CCM.0000000000004916.

19. Koyner, J.L.; Carey, K.A.; Edelson, D.P.; Churpek, M.M. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit Care Med* **2018**, *46*, 1070-1077, doi:10.1097/CCM.0000000000003123.

**Footnotes**

Contributions: Conceptualization, QL and MH; Methodology, YX; Software, YX; Validation, YZ, and YX.; Formal Analysis, JC and XC; Resources, MH; Data Curation, QL; Writing – Original Draft Preparation, QL; Writing – Review & Editing, MH; Visualization, YX; Supervision, YZ and QL; Project Administration, QL;

**Funding**: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest: Not Applicable.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Acknowledgements: Thank my colleagues in general ICU for cooperating with us in our prospective study. Thanks to two public databases namely MIMIC and AmsterdamUMC for providing important data.

**Supplemental Figure:**

File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.

File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.

File name: **Figure S3. tif**

Title of data: The visual interface for prospective studies of the AKI early diagnosis prediction models.

Description of data: Different patients have independent card display information, in which the last line of the card is the AKI diagnosis. The visual interface was not available to doctors during the prospective study and was used only by the researchers to verify the diagnosis and for sampling verification.

File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

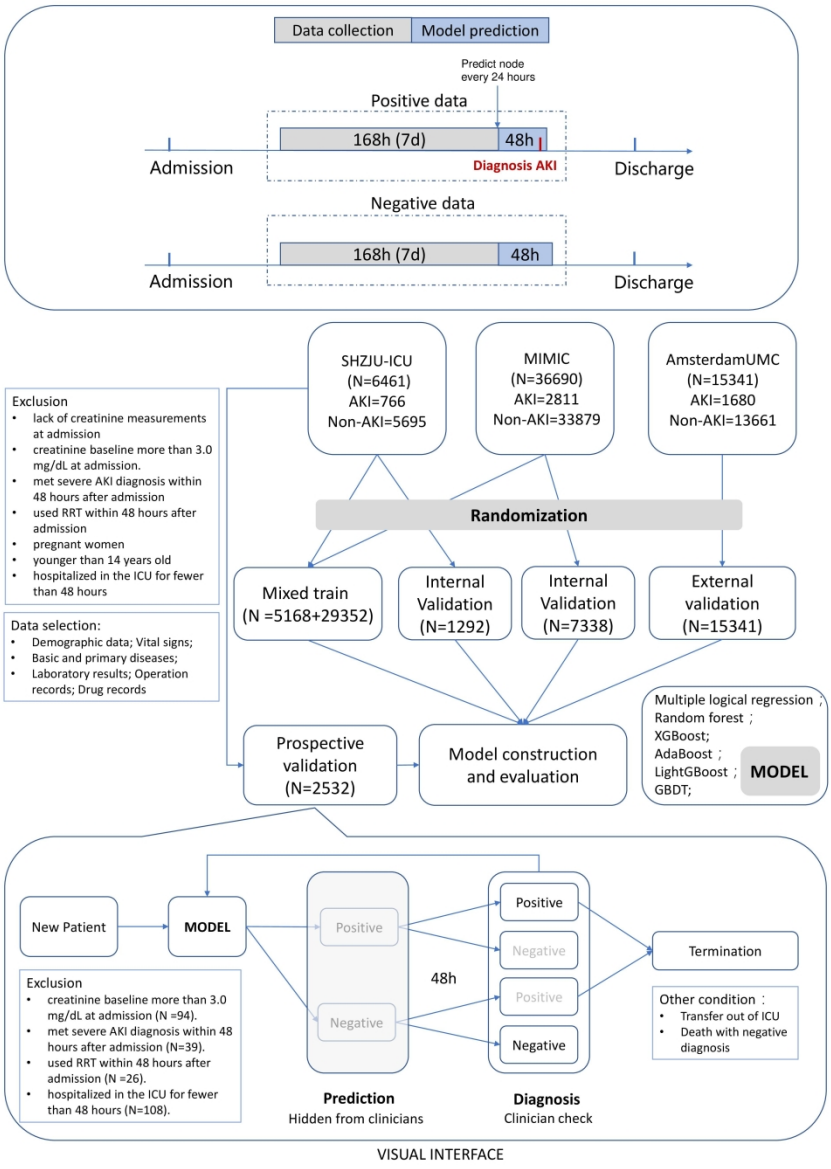
Description of data: Writing norms of this type of research.

**Figure Legend:**

Figure 1: The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

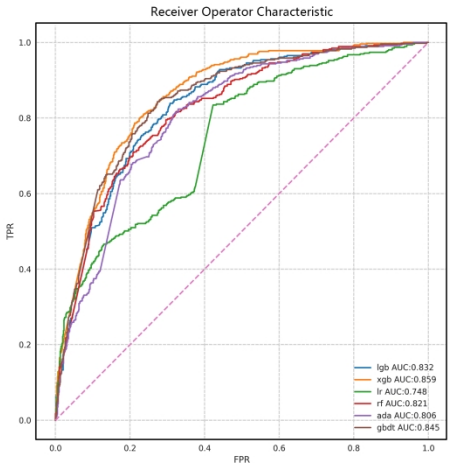
Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.



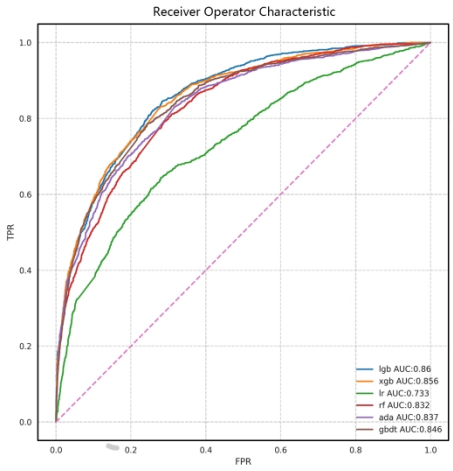

338x451mm (300 x 300 DPI)

Internal validation set ROC curves

A: SHZJU-ICU database validation set

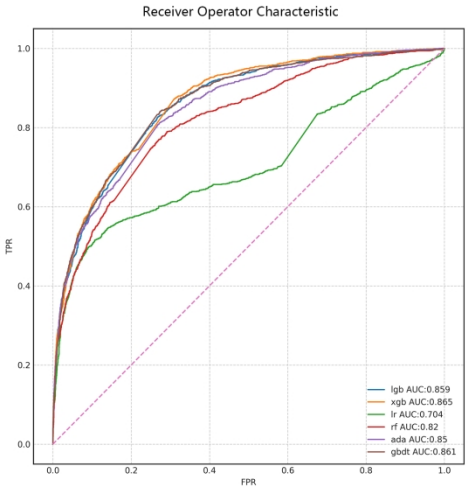



B: MIMIC database validation set

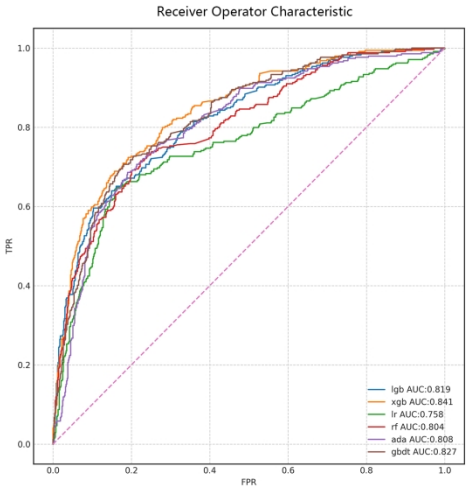


External and prospective validation set ROC curves

A: AmsterdamUMC database external validation set



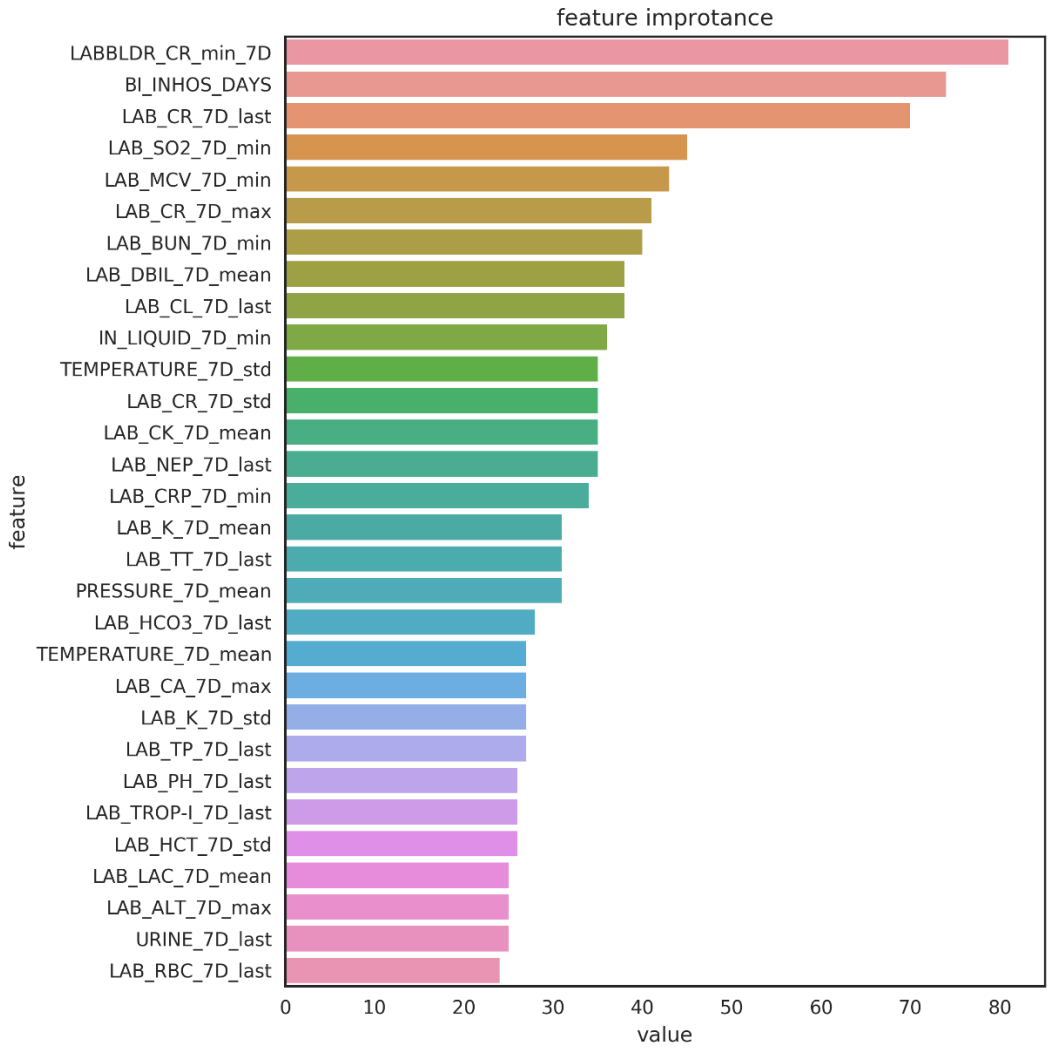

B: SHZJU-ICU database prospective validation set



**Supplemental Figure:**

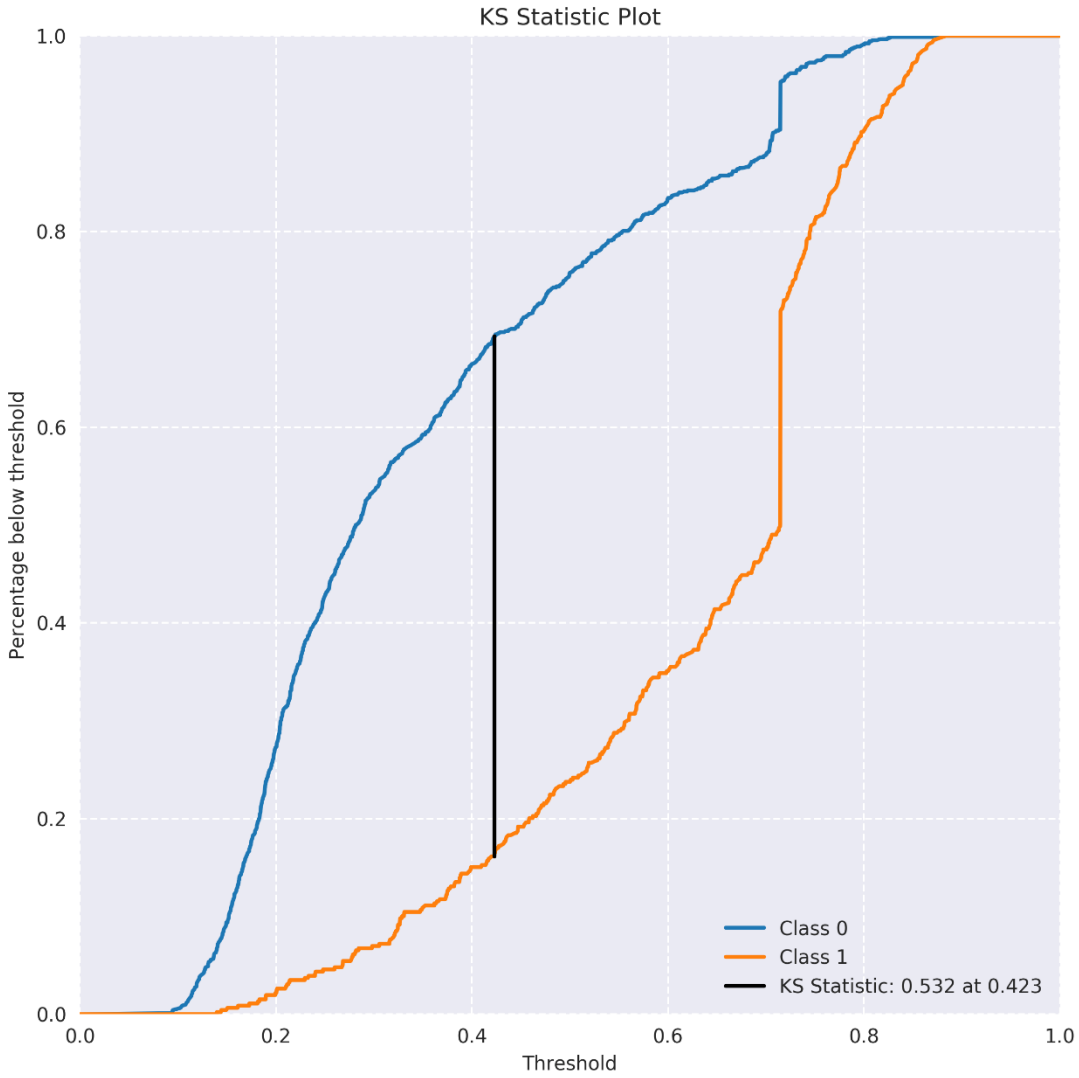
File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.




File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.




File name: **Figure S3. tif**

Title of data: The visual interface for prospective studies of the AKI early diagnosis prediction models.

Description of data: Different patients have independent card display information, in which the last line of the card is the AKI diagnosis. The visual interface was not available to doctors during the prospective study and was used only by the researchers to verify the diagnosis and for sampling verification.


File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

| | SHZJU-ICU | MIMIC | AmsterdamUMC | Including in Model |
|---|---|---|---|---|
| **Age** | √ | √ | √ | √ |
| **Gender** | √ | √ | √ | √ |
| **Race** | √ | √ | × | × |
| **Primary disease** | √ | √ | × | × |
| **Comorbidity** | √ | √ | × | × |
| **Ventilation** | √ | √ | √ | √ |
| **Operation** | √ | √ | × | × |
| **ICU hours** | √ | √ | √ | √ |
| **Survived** | √ | √ | √ | √ |
| **Vital sign** | | | | |
| **Temperature** | √ | √ | √ | √ |
| **Systolic pressure** | √ | √ | √ | √ |
| **Diastolic pressure** | √ | √ | √ | √ |
| **Respiratory rate** | √ | √ | √ | √ |
| **Heart rate** | √ | √ | √ | √ |
| **Oxygen saturation** | √ | √ | √ | √ |
| **Urine** | √ | √ | √ | √ |
| **GCS score** | √ | × | × | × |
| **Laboratory results** | | | | |

| | | | | |
|---|---|---|---|---|
| **Scr** | √ | √ | √ | √ |
| **BUN** | √ | √ | √ | √ |
| **RBC** | √ | √ | √ | √ |
| **MCV** | √ | √ | √ | √ |
| **Hb** | √ | √ | √ | √ |
| **WBC** | √ | √ | √ | √ |
| **NEUT** | √ | √ | √ | √ |
| **PLT** | √ | √ | √ | √ |
| **HCV** | √ | √ | × | × |
| **TBLB** | √ | √ | √ | √ |
| **DBLB** | √ | √ | √ | √ |
| **IBLB** | √ | √ | √ | √ |
| **CRP** | √ | √ | √ | √ |
| **PCT** | √ | × | × | × |
| **Serum kalium** | √ | √ | √ | √ |
| **Serum natrium** | √ | √ | √ | √ |
| **Serum chlorine** | √ | √ | √ | √ |
| **PT** | √ | √ | √ | √ |
| **APTT** | √ | √ | √ | √ |
| **INR** | √ | √ | √ | √ |
| **CK** | √ | √ | √ | √ |
| **LDH** | √ | √ | √ | √ |
| **Troponin** | √ | √ | √ | √ |
| **Blood glucose** | √ | √ | √ | √ |
| **PH** | √ | √ | √ | √ |
| **Lactic acid** | √ | √ | √ | √ |
| **Anion gap** | √ | √ | √ | √ |
| **ABC** | √ | √ | √ | √ |
| **SBC** | √ | √ | √ | √ |
| **$PaCO_2$** | √ | √ | √ | √ |
| **$PaO_2$** | √ | √ | √ | √ |

SHZJU-ICU: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; MIMIC: Medical Information Mart for Intensive Care; APTT: Activated partial thromboplastin time; INR: International normalized ratio; LDH: Lactate dehydrogenase; $PaCO_2$: Partial pressure of carbon dioxide; $PaO_2$ Partial pressure of oxygen; MCV: Mean corpuscular volume; PH: Potential of hydrogen; WBC: White blood cell count; NEUT: Proportion of neutrophils; CK: Creatine kinase; Scr: Serum Creatinine; BUN: Blood urea nitrogen; RBC: Red blood cell count; Hb: Hemoglobin; PT: Prothrombin time; TBLB: Total bilirubin; DBLB: Direct bilirubin; IBLB: Indirect bilirubin; CRP: C-reactive protein; HCV: Hematocrit value; PCT: Procalcitonin; ABC: Actual Bicarbonate Radical; SBC: Standard Bicarbonate Radical; PLT: Platelet count

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | SHZJU-ICU Prospective Negative (N=2188) | SHZJU-ICU Prospective Positive (N=344) | P-value |
|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 61.3 [47.5-73.4] | 62.1 [51.2-70.5] | 0.03 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 1365 (62.4) | 207 (60.2) | 0.438 |
| **Race** | | | | | | |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 2 (0.001) | 0 | 1 |
| **Black** | 1 (0.02) | 0 | 1 | 0 | 0 | 1 |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 2185 (99.9) | 344(100) | 1 |
| **Other** | 0 | 1 (0.1) | 1 | 1 (0.001) | 0 | 1 |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 0.8 (0.32) | 0.9 (0.35) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (11) | 20 (10) | 0.007 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 230 (10.5) | 26 (7.6) | 0.102 |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 15 (0.7) | 3 (0.8) | 0.726 |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 98 (4.5) | 19 (5.7) | 0.403 |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 142 (6.5) | 29 (8.6) | 0.203 |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 324(14.8) | 56 (16.4) | 0.466 |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 1102 (50.4) | 239 (69.5) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 1234(56.4) | 160 (46.7) | 0.001 |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 58.7 [34.2-160.4] | 145.6 [68.5-314.8] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 1868 (85.4) | 178 (51.9) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

| 1Section/Topic | | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L80 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L84-86 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P5, L104 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L116 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L119 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L126 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P6, L159 |
| | 10e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | NA |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |

| | | | | |
|---|---|---|---|---|
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 3a | ;V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 3b | ;V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 |
| | 3c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, Table2 |
| Model development | 4a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 4b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes |
| Model specification | 5a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 5b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | ;V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | ;V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L236 |
| Interpretation | 9a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L217 |
| | 9b | ;V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L198 |
| Implications | 20 | ;V | Discuss the potential clinical use of the model and implications for future research. | Yes, P9, L193 |
| **Other information** | | | | |
| Supplementary information | 21 | ;V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | ;V | Give the source of funding and the role of the funders for the present study. | NA |

BMJ Open

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center perspective study



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2021-054092.R2 |
| Article Type: | Original research |
| Date Submitted by the Author: | 21-Nov-2021 |
| Complete List of Authors: | Liang, Qiqiang; Zhejiang University School of Medicine Second Affiliated Hospital<br>Xu, Yongfeng; HealSci Technology Co<br>Zhou, Yu; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Xinyi; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Juan; Zhejiang University School of Medicine Second Affiliated Hospital<br>Huang, Man; Zhejiang University School of Medicine Second Affiliated Hospital, |
| Primary Subject Heading: | Intensive care |
| Secondary Subject Heading: | Diagnostics |
| Keywords: | Acute renal failure < NEPHROLOGY, INTENSIVE & CRITICAL CARE, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center perspective study

Qiqiang Liang[1]; Yongfeng Xu[2]; Yu Zhou[1]; Xinyi Chen[1], MM; Juan Chen[1]; Man Huang[1*]

[1]General Intensive Care Unit, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

[2]BSc, HealSci Technology Co., Ltd, Beijing China.

*Corresponding author: Man Huang, MD PhD.

Department: General Intensive Care Unit

Hospital: Second Affiliated Hospital of Zhejiang University School of Medicine

Address: No. 1511, Jianghong Road, Bingjiang District, Hangzhou City, Zhejiang Province

Tel: + 86 571 89713427

Fax: + 86 571 89713427

E-mail: huangman@zju.edu.cn.

ORCID: 0000-0002-0842-2254

**Abstract:**

Objectives: There are many studies of acute kidney injury (AKI) diagnosis models lack of external validation and prospective validation. We constructed the models using three databases to predict severe AKI within 48 hours in intensive care unit (ICU) patients.

Design: A retrospective and prospective cohort study.

Setting: We studied critically ill patients in our database (SHZJU-ICU) and two other public databases, the Medical Information Mart for Intensive Care (MIMIC) and AmsterdamUMC databases, including basic demographics, vital signs, and laboratory results. We predicted the diagnosis of severe AKI in patients in the next 48 hours using machine-learning algorithms with the three databases. Then, we carried out real-time severe AKI prediction in the prospective validation study at our centre for one year.

Participants: All patients included in three databases with uniform exclusion criteria.

Primary and secondary outcome measures: Effect evaluation index of prediction models.

Results: We included 58492 patients, and a total of 5257 (9.0%) patients met the definition of severe AKI. In the internal validation of the SHZJU-ICU and MIMIC databases, the best area under the receiver operating characteristic curve (AUROC) of the model was 0.86. The external validation results by AmsterdamUMC database were also satisfactory, with the best AUROC of 0.86. A total of 2532 patients were admitted to the center for prospective validation; 358 positive results were predicted, and 344 patients were diagnosed with severe AKI, with the best sensitivity of 0.72, the specificity of 0.80, and the AUROC of 0.84.

Conclusion: The prediction model of severe AKI exhibits promises as a clinical application based on dynamic vital signs and laboratory results of multi-center databases with prospective and external validation.

**Keywords**: Machine learning; Acute kidney injury; Real-time prospective validation; External validation.

**Strengths and limitations of this study**:

A prospective validation in machine learning of AKI research rather than other studies;

Three large database containing different national populations and regions;

Variable's sampling limited by the monitoring frequency of clinical data;

Differences in the samples proportion of three databases;

The dimensions of variables are not rich enough.

## Introduction:

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalization and mortality[1]. AKI is divided into three types according to the various aetiologies: prerenal (renal hypoperfusion), intrarenal (vascular, glomerular, or tubulointerstitial lesions), and postrenal (urinary tract obstruction)[2]. Although nearly all diseases associated with ICU admission may cause AKI, acute tubular necrosis (ATN) and prerenal azotaemia are the most common causes[3].

All AKI diagnostic criteria including the latest Kidney Disease: Improving Global Outcomes (KDIGO) standard are currently based on the creatinine level and urine volume[4]. However, the increase in the creatinine level or decrease in the urine volume lags the onset of AKI[2]. Many studies have suggested that early diagnosis and treatment of reversible AKI can reduce mortality[5]. Therefore, the creatinine level and urine volume are not satisfactory to meet clinical diagnostic demands. Consequently, many researchers have tried to develop an early warning model by analysing the risk factors for AKI[6].

Patient complications, such as diabetes, hypertension, cardiovascular disease, chronic liver disease, sepsis, and trauma, are identified as important risk factors for AKI[7]. The AKI prediction model and scoring system developed based on high-risk factors has gradually become the focus of research considering the lower clinical application threshold compared with that of new biomarkers[6]. Although most previous prediction models use the multiple logistic regression model, a variety of AKI prediction models based on machine learning have resulted in satisfactory outcomes[6]. Since the first AKI prediction model study based on artificial intelligence was published in 2016, researchers have built more than 20 published AKI prediction models successively by using local or multi-centre databases[6,8-14]. The results indicate that these models can predict the occurrence of AKI and the need for renal replacement therapy (RRT) within 24 or 48 hours, with accuracies ranging from 81% to 97%[6,15]. In addition, many studies have focused on subspecialized conditions, including cardiac surgery, trauma, and burns[14-16]. However, the common defect in these studies is the lack of external validation and prospective validation, which causes the prediction model to deviate from the clinical scenarios and limits extrapolation beyond the scope of the data.

In this study, we built models to predict AKI within 48 hours in critically ill patients by using transcontinental three databases. Then, we evaluated the clinical effect of the model through a one-year prospective validation at our centre.

## Methods:

### Study design and setting

We collected patients using three ICU databases and prospectively validated the models in our centre. The first database was our centre general ICU database (SHZJU-ICU) of the Second Affiliated Hospital

of Zhejiang University School of Medicine, an academic teaching hospital. Since its establishment in 2017, it has included 12000 ICU patients' data and is updated daily. The Medical Information Mart for Intensive Care (MIMIC) III database, the second one, is an open ICU database provided by the Massachusetts Institute of Technology and includes nearly 60000 ICU patients from North America[17]. Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23000 patients admitted to ICUs in parts of Europe[18]. The research flow chart is shown in Figure 1.

**Study definition**

In this study, the diagnosis of AKI was confirmed based on three stages according to the KDIGO criteria[4]. We defined the patients who met the KDIGO AKI II and III criteria as severe AKI groups and the others as negative groups. We excluded patients with lack of creatinine measurements during admission, patients with creatinine baseline more than 3.0 mg/dL at admission, patients who met severe AKI diagnosis within 24 hours, and patients who used RRT within 48 hours after admission[19]. In addition, we excluded pregnant women, patients younger than 14 years old, and patients hospitalized in the ICU for fewer than 48 hours. After the patient was admitted to ICU, we performed a prediction every 24 hours and recorded a prediction time. If the patient was diagnosed with severe AKI within 48 hours, the predictive time was defined as a positive predictive point, and the others were defined as a negative point. The study was evaluated and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine as study number 2019-078.

**Data Collection**

The variables included demographic data, vital signs, basic and primary diseases, laboratory results, important operation records, and drug records. Comorbidity included hypertension, diabetes, cardiopathy,

liver disease, and malignant tumors. The primary disease was the main cause of admission to the ICU following the ICD-10 codes. The vital signs and clinical laboratory results were transformed into different variables according to the average, variance, maximum, minimum, and final value before diagnosis. We use a method similar to the forward incremental method in the multivariate logic regression model, that is, the combination of embedded feature selection and forward addition for feature selection. First of all, all variables are trained in the model, then list by variables importance. variables are added to the model one by one according to the variable importance. a variable is retained if it causes the AUC growth to be greater than 0.01, otherwise delete it. We transformed the MIMIC and AmsterdamUMC databases according to our centre database structure, unifying the unit and diagnostic codes. We deleted variables missing more than 50%. Variables missing more than 30% but less than 50% are listed to clinicians who determine the potential correlation between these variables and AKI. We carry out multiple interpolation for these variables which clinicians require to be retained, and the others deleted. Variables missing less than 30% are fill in multiple interpolation. All Missing data between three databases and values included in the model shown in supplementation file Table S1.

**Model construction and external validation**

The ratio of the training and internal validation sets was 4:1. The SHZJU and MIMIC databases training sets were mixed into a new training set. There were more negative data than positive data, so we randomly sampled the negative datasets and constructed a new data subset with a sampling ratio of positive and negative data of 1:5 in model building in order to extract the importance variables. In the subsequent model validation, we adopted the original data set. We used multiple logistic regression, random forest, XGBoost, AdaBoost, LightGBoost, gradient boosting decision tree (GBDT), and debug to assess the variables and model-related parameters by the fivefold cross-validation method. After the

models were built, we used the SHZJU-ICU and MIMIC test sets for internal validation of the model and the AmsterdamUMC database for external validation. The most appropriate cut-off value was determined according to the K-S curve. The prediction model represents the results of each prediction with a probability between 0 to 1.0. We define results more than 0.4 as high-risk, that is, positive results, and the rest as negative results. Through internal validation and external validation, we calibrated the model by adjusting the super-parameters and using the Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. All model building and validation processes were performed in Python 3.6.

**Prospective validation**

The prospective research period was 2020.01.01 to 2020.12.31. We collected real-time data when patients were admitted to the ICU, transformed the data according to the requirements, and formed a complete sample for the prediction model after passing the integrity test. We had established a visualization scheme and allowed researchers to review the predictions daily. The daily prediction results were not publicly accessible during the study to avoid affecting clinicians' decisions, but the diagnosis results were available to the researchers as visual graphics. We sampled the 20% predicted data every month and deleted samples with more than 50% missing values to ensure data correctness. When a patient has the following conditions, AKI prediction system will end the patient's prospective prediction: A. a positive diagnosis; B. Transfer out of ICU or death with negative diagnosis. All diagnosis of severe AKI needs to be reviewed by two ICU attending physicians independently, and if the they have different opinion, the third one will be appealed.

**Statistical analysis:**

The population characteristics were reported as the medians and inter-quartile ranges (IQRs) for

skewed data and the means and standard deviations for normally distributed data. The independent sample T-test was used for normally distributed data, and the rank-sum test was used for the rest. Dichotomous variables were assessed by the chi-square test, and a P value less than 0.05 was considered statistically significant. The non-normally distributed data were analyzed by exponential transformation and logarithmic transformation. The effect of the model was evaluated by parameters such as the AUROC, accuracy, specificity, and F1-score.

**Patient and Public Involvement:**

The information of cases in three databases was in a state of complete desensitization in the process of building the model. During the prospective study, all the patients signed an informed consent form at the beginning of admission to ICU. The real-time data discussed and used by only the study members, and were not made public during the study period. All data were anonymized before the authors accessed them for the purpose of this study. Therefore, patients' priorities, experience, and preferences will not affect the development of the research question and outcome measures. If necessary, we will inform patients of relevant research results by telephone.

**Results:**

According to the inclusion criteria and exclusion criteria, we selected 58492 patients from three databases who met the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36690 patients from the MIMIC database, and 15341 patients from the AmsterdamUMC database. A total of 5257 (9.0%) patients met the definition of severe AKI (11.8% in SHZJU-ICU, 7.6% in MIMIC, and 10.9% in AmsterdamUMC). The distributions of age and sex in the three centers were similar, but the differences in race were large. Asian patients accounted for more than 99% in the SHZJU-ICU

database, and only approximately 2.5% in the MIMIC database. White people accounted for more than 70% of the MIMIC database. In addition, patients from the MIMIC database had a higher incidence of the tumor, liver cirrhosis, diabetes, and hypertension. Patients in the AmsterdamUMC and SHZJU-ICU databases had a higher proportion of mechanical ventilation and overall survival rate. Severe AKI patients had longer ICU hospital stays and higher mortality. More details are presented in Table 1.

Table 1: Clinical Demographics and Outcomes in patients with or without severe acute kidney injury.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | MIMIC Negative (N=33879) | MIMIC Severe AKI (N=2811) | P-value | AmsterdamUMC Negative (N=13661) | AmsterdamUMC severe AKI (N=1680) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 62.0 [45.0-76.0] | 64.0 [52.0-76.0] | <0.001 | 64.5 [54.5,74.5] | 64.3 [52.5,78.5] | 0.04 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 18976 (56.0) | 1524 (54.2) | 0.06 | 8920 (65.3) | 1054 (63.9) | 0.03 |
| **Race** | | | | | | | / | / | / |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 23796 (70.2) | 2005 (71.3) | 0.23 | / | / | / |
| **Black** | 1 (0.02) | 0 | 1 | 2455 (7.2) | 199 (7.1) | 0.45 | / | / | / |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 918 (2.7) | 70 (2.5) | 0.54 | / | / | / |
| **Other** | 0 | 1 (0.1) | 1 | 6710 (19.8) | 537 (19.1) | 0.34 | / | / | / |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 1.1 (0.41) | 1.1 (0.50) | 0.04 | 1.2 (0.35) | 1.1 (0.38) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (12) | 23 (12) | <0.001 | 18 (13) | 23 (10) | <0.001 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 6652 (19.6) | 752 (26.8) | <0.001 | / | / | / |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 1180 (3.5) | 426 (15.2) | <0.001 | / | / | / |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 4840 (14.3) | 336 (12.0) | 0.001 | / | / | / |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 7250 (21.4) | 789 (28.1) | <0.001 | / | / | / |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 16328 (48.2) | 1513 (53.8) | <0.001 | / | / | / |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 13009 (38.4) | 1503(53.5) | <0.001 | 7718(56.5) | 1248(74.3) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 12455 (36.7) | 1430 (50.8) | <0.001 | / | / | / |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 51.0 [28.5-110.3] | 103.4 [50.1-261.3] | <0.001 | 24.0 [19.8,63.5] | 142.0 [45.4,394.2] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 25648 (75.7) | 1356 (48.2) | <0.001 | 10942(80.1) | 1033 (61.5) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

There were significant differences in the important parameters of the variables among the different models (see Figure S1). However, the trend of the creatinine level in the past week was still an important variable, followed by urine volume, blood urea nitrogen level, temperature, and length of ICU stay. The cut-off value used to distinguish between a negative and positive prediction was determined by the K-S curve, with value of 0.423 (see Figure S2). The GBDT model had the best prediction effect in the test set, followed by XGBoost and LightGBoost. In the two central internal validation sets, the two best-performing machine learning algorithms with great AUROC are LightGBoost (SHZJU-ICU of 83.2%, MIMIC of 86.0%) and XGBoost (SHZJU-ICU 85.9%, MIMIC 85.6%), as detailed in Figure 2. Overall, the sensitivity (SHZJU-ICU 0.84, MIMIC 0.83) and the negative predictive value (SHZJU-ICU 0.90, MIMIC 0.90) of the predictive model were high, but the specificity was general (SHZJU-ICU 0.79, MIMIC 0.75), as shown in Table 2. In the external validation based on AmsterdamUMC database, the overall model validation effect results were satisfactory, and XGBoost had the best performance, with an AUROC of 0.84, as shown in Figure 2 and Table 2. We built a visual prediction interface based on the prediction model (supplementation file Figure S3).

Table 2: Model validation results by three databases with machine learning algorithm

| *Model* | *AUROC* | *Accuracy* | *Sensitivity* | *Specificity* | *PPV* | *NPV* | *F1* |
|---|---|---|---|---|---|---|---|
| ***Internal validation with SHZJU-ICU database*** | | | | | | | |
| ***Logistic regression*** | *0.748* | *0.662* | *0.834* | *0.576* | *0.496* | *0.874* | *0.622* |
| ***LightGBoost*** | *0.832* | *0.741* | *0.839* | *0.692* | *0.576* | *0.896* | *0.683* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ***GBDT*** | *0.845* | *0.765* | *0.843* | *0.725* | *0.606* | *0.902* | *0.705* |
| ***AdaBoost*** | *0.806* | *0.721* | *0.824* | *0.67* | *0.555* | *0.884* | *0.663* |
| ***Random Forest*** | *0.821* | *0.763* | *0.71* | *0.789* | *0.627* | *0.845* | *0.666* |
| ***XGBoost*** | ***0.859*** | *0.779* | *0.81* | *0.763* | *0.631* | *0.889* | *0.709* |
| ***Internal validation with MIMIC database*** | | | | | | | |
| ***Logistic regression*** | *0.733* | *0.695* | *0.643* | *0.72* | *0.535* | *0.801* | *0.584* |
| ***LightGBoost*** | ***0.86*** | *0.768* | *0.822* | *0.741* | *0.613* | *0.893* | *0.702* |
| ***GBDT*** | *0.846* | *0.765* | *0.786* | *0.755* | *0.616* | *0.876* | *0.691* |
| ***AdaBoost*** | *0.837* | *0.732* | *0.831* | *0.683* | *0.567* | *0.89* | *0.674* |
| ***Random Forest*** | *0.832* | *0.738* | *0.791* | *0.712* | *0.578* | *0.872* | *0.668* |
| ***XGBoost*** | *0.856* | *0.758* | *0.833* | *0.721* | *0.598* | *0.895* | *0.695* |
| ***External validation with AmsterdamUMC database*** | | | | | | | |
| ***Logistic regression*** | *0.704* | *0.767* | *0.516* | *0.893* | *0.706* | *0.787* | *0.596* |
| ***LightGBoost*** | *0.859* | *0.763* | *0.827* | *0.731* | *0.606* | *0.894* | *0.7* |
| ***GBDT*** | *0.861* | *0.764* | *0.84* | *0.727* | *0.606* | *0.901* | *0.704* |
| ***AdaBoost*** | *0.85* | *0.755* | *0.813* | *0.726* | *0.597* | *0.886* | *0.689* |
| ***Random Forest*** | *0.82* | *0.743* | *0.77* | *0.729* | *0.587* | *0.864* | *0.666* |
| ***XGBoost*** | ***0.865*** | *0.75* | *0.873* | *0.688* | *0.584* | *0.916* | *0.7* |
| ***Prospective validation with SHZJU-ICU*** | | | | | | | |
| ***Logistic regression*** | *0.758* | *0.772* | *0.648* | *0.834* | *0.662* | *0.826* | *0.655* |
| ***LightGBoost*** | *0.819* | *0.796* | *0.596* | *0.895* | *0.74* | *0.816* | *0.66* |
| ***GBDT*** | *0.827* | *0.781* | *0.706* | *0.818* | *0.66* | *0.848* | *0.683* |
| ***AdaBoost*** | *0.808* | *0.766* | *0.686* | *0.805* | *0.638* | *0.837* | *0.661* |
| ***Random Forest*** | *0.804* | *0.755* | *0.715* | *0.775* | *0.613* | *0.845* | *0.66* |
| ***XGBoost*** | ***0.841*** | *0.779* | *0.724* | *0.807* | *0.652* | *0.854* | *0.686* |

*AUROC*: Area Under Receiver Operating Characteristic; *PPV*: Positive Predictive Value; *NPV*: Negative Predictive Value; *SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *GBDT*: Gradient Boosted Decision Tree.

According to the inclusion and exclusion criteria, we delete 267 patients among 94 patients with creatinine baseline more than 3.0 mg/dL at admission, 39 patients met severe AKI diagnosis within 24 hours, and 26 patients who used RRT within 48 hours after admission, 108 patients hospitalized in the ICU for fewer than 48 hours. A total of 2532 patients were admitted to our centre for prospective validation, and the prediction model made 16858 times predictions. In the prospective cohort, there was no significant difference in age, gender, baseline creatinine and urea nitrogen, and complications. The

proportion of mechanical ventilation and the ICU stay time in AKI patients were longer with higher mortality. Above all, there was no significant difference between the prospective and the retrospective cohort. More detail sees in supplementation file Table S2. In the end, 358 positive results were predicted, and the rest were negative results. There are 344 patients with severe AKI were diagnosed and the prediction accuracy was 83.5%. The model with the highest area under the curve was XGBoost, 0.84 with the best sensitivity of 0.72, the specificity of 0.80. The results of the prospective study are similar to those of the external validation of the model, and are relatively stable. More detail is presented in Figure 3 and Table 2.

**Discussion:**

In this study, we built predictive models by machine learning to predict the incidence of severe AKI with three databases in different regions and in the next 48 hours. After internal and external validation, prospective validation over one year was carried out to verify the model effects. The three databases come from three countries that are in Asia, Europe, and North America, which proves that the model is universal to some extent.

Despite the huge amount of data, many databases are still not suitable for prospective research because they are not updated promptly. Tomasev, N and colleagues have provided research on AKI prediction models with a large amount of data[10]. The study covered 703782 adult patients with 6 billion individual items, including 620000 elements. In this study, a depth neural network model was used for real-time prediction. A total of 55.8% of severe AKI patients were predicted within the first 48 hours, although each accurate prediction was accompanied by two mispredictions[10]. This study provided a new scheme for real-time prediction and indicated that we should prospectively evaluate and

independently validate models to explore their effectiveness. In a prospective study, Flechet, M et al compared an AKI prediction model with clinicians in 252 patients and found that the clinical effect of the random forest model for predicting AKI-II/III was equivalent to that of clinicians. Our prediction model graphical visualization of the model was installed in the centre's database for better usage. In addition, our database is updated daily to achieve daily predictions and present the results to researchers. In the prospective validation of our study, the stability of the prediction model confirmed its promise, which provides a basis for future research.

There are many studies of artificial intelligence for predicting the occurrence of AKI, but most of them are single-centre studies, and the extrapolation effect has been controversial. In 2016, Koyner, J, and his colleague published the first study of an AKI prediction model based on multi-centre data. In all 202961 patients, 17541 (8.6%) had AKI, 4251 (3.5%) had AKI-II, and 1242 (0.6%) had AKI-III. A multivariate logistic regression model was used to predict AKI in this study with an AUROC of 0.74 (95% CI=0.74). With the classification of AKI, the AUROC of the prediction model gradually increased to 0.84[8]. Subsequently, the study team used a new machine learning algorithm, to build a more accurate model to predict the occurrence of AKI-II, with an AUROC of 0.9 within 24 hours and 0.87 within 48 hours[19]. Recently, the research team included data from two other centres, namely, LUMC (N=200613) and NUS (N=246895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85 to 0.86, suggesting that the artificial intelligence model has stable predictive ability[12]. This series of studies included many data points, suggesting the feasibility of artificial intelligence in the diagnosis of AKI, but the proportion of positive patients (3.5%) and ICU patients (30%) was too low to properly predict AKI. Our research is similar to the above. The SAHZJU-ICU database is a single-centre database representing south-eastern China, and the MIMIC database is a well-known open ICU database in the

United States. The AmsterdamUMC database is a public database located in Europe. The population structure and diseases in the three databases are complete but different in the distribution of complications and race. Therefore, it provides a prediction model with unparalleled stability compared with other studies.

Limitations:

This retrospective multi-centre study was unable to carry out more clinical feature mining and comparison because of different data structures. The differences between the three databases partly reflect some demographic differences between Europe, the United States and China, resulting in a decline in the accuracy of the prediction model. There are some differences in the number of patients included in the three databases, which may affect the choice of variables. As a result of the study design, we deleted patients with ICU hospitalization of less than 48 hours, which may result in the exclusion of most relatively mild patients and may reduce false positives. Second, in the prospective data study in 2020, there may be deviations in the inclusion of patients in the centre, thus affecting the interpretation of the follow-up prospective results. Finally, given the low incidence of severe AKI and the great difference in the proportion of positive and negative samples, the data may be accidental. Our model seems to be superior to diagnostic non-AKI patients rather than AKI because of the proportion of positive data that we include. In the retrospective study, we reduced the proportion of negative data by randomization but retain all date in prospective phase with the sensitivity decreases.

**Conclusion:**

Based on databases of patients of different races from different countries, we constructed stable

machine learning models to predict the occurrence of AKI in the next 48 hours. Prospective validation through the implementation of an updated local database is an effective exploration of further research.

Data Availability Statement:

The model code can be obtained by email if readers need it, but we cannot guarantee that all the code will be provided. Two public databases can be applied from the official website of their respective databases.

Ethics approval:

Our research was approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

Competing interests for authors:

The authors declare that they have no competing interests.

**Reference:**

1. Ronco, C.; Bellomo, R.; Kellum, J.A. Acute kidney injury. *The Lancet* **2019**, *394*, 1949-1964, doi:10.1016/s0140-6736(19)32563-2.
2. Pakula, A.M.; Skinner, R.A. Acute Kidney Injury in the Critically Ill Patient: A Current Review of the Literature. *J Intensive Care Med* **2016**, *31*, 319-324, doi:10.1177/0885066615575699.
3. Bellomo, R.; Kellum, J.A.; Ronco, C. Acute kidney injury. *The Lancet* **2012**, *380*, 756-766, doi:10.1016/s0140-6736(11)61454-2.
4. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin*

*Pract* **2012**, *120*, c179-184, doi:10.1159/000339789.

5. Coelho, S.; Cabral, G.; Lopes, J.A.; Jacinto, A. Renal regeneration after acute kidney injury. *Nephrology (Carlton)* **2018**, *23*, 805-814, doi:10.1111/nep.13256.

6. Wilson, F.P.; Greenberg, J.H. Acute Kidney Injury in Real Time: Prediction, Alerts, and Clinical Decision Support. *Nephron* **2018**, *140*, 116-119, doi:10.1159/000492064.

7. Wu, Y.; Hao, W.; Chen, Y.; Chen, S.; Liu, W.; Yu, F.; Hu, W.; Liang, X. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* **2020**, *42*, 1127-1134, doi:10.1080/0886022X.2020.1843491.

8. Koyner, J.L.; Adhikari, R.; Edelson, D.P.; Churpek, M.M. Development of a Multicenter Ward-Based AKI Prediction Model. *Clin J Am Soc Nephrol* **2016**, *11*, 1935-1943, doi:10.2215/CJN.00280116.

9. Low, S.; Vathsala, A.; Murali, T.M.; Pang, L.; MacLaren, G.; Ng, W.Y.; Haroon, S.; Mukhopadhyay, A.; Lim, S.L.; Tan, B.H., et al. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* **2019**, *20*, 32, doi:10.1186/s12882-019-1206-4.

10. Tomasev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I., et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116-119, doi:10.1038/s41586-019-1390-1.

11. Flechet, M.; Falini, S.; Bonetti, C.; Güiza, F.; Schetz, M.; Van den Berghe, G.; Meyfroidt, G. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Critical Care* **2019**, *23*, 282,

doi:10.1186/s13054-019-2563-x.

12. Churpek, M.M.; Carey, K.A.; Edelson, D.P.; Singh, T.; Astor, B.C.; Gilbert, E.R.; Winslow, C.; Shah, N.; Afshar, M.; Koyner, J.L. Internal and External Validation of a Machine Learning Risk Score for Acute Kidney Injury. *JAMA Netw Open* **2020**, *3*, e2012892, doi:10.1001/jamanetworkopen.2020.12892.

13. Yuan, Q.; Zhang, H.; Deng, T.; Tang, S.; Yuan, X.; Tang, W.; Xie, Y.; Ge, H.; Wang, X.; Zhou, Q., et al. Role of Artificial Intelligence in Kidney Disease. *Int J Med Sci* **2020**, *17*, 970-984, doi:10.7150/ijms.42078.

14. Li, Y.; Xu, J.; Wang, Y.; Zhang, Y.; Jiang, W.; Shen, B.; Ding, X. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* **2020**, *43*, 752-761, doi:10.1002/clc.23377.

15. De Vlieger, G.; Kashani, K.; Meyfroidt, G. Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review. *Curr Opin Crit Care* **2020**, *26*, 563-573, doi:10.1097/MCC.0000000000000775.

16. James, M.T.; Pannu, N.; Hemmelgarn, B.R.; Austin, P.C.; Tan, Z.; McArthur, E.; Manns, B.J.; Tonelli, M.; Wald, R.; Quinn, R.R., et al. Derivation and External Validation of Prediction Models for Advanced Chronic Kidney Disease Following Acute Kidney Injury. *JAMA* **2017**, *318*, 1787-1797, doi:10.1001/jama.2017.16326.

17. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.W.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.A.-O.; Celi, L.A.; Mark, R.A.-O. MIMIC-III, a freely accessible critical care database.

18. Thoral, P.J.; Peppink, J.M.; Driessen, R.H.; Sijbrands, E.J.G.; Kompanje, E.J.O.; Kaplan, L.; Bailey, H.; Kesecioglu, J.; Cecconi, M.; Churpek, M., et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. *Crit Care Med* **2021**, 10.1097/CCM.0000000000004916, doi:10.1097/CCM.0000000000004916.

19. Koyner, J.L.; Carey, K.A.; Edelson, D.P.; Churpek, M.M. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit Care Med* **2018**, *46*, 1070-1077, doi:10.1097/CCM.0000000000003123.

**Footnotes**

Contributions: Conceptualization, QL and MH; Methodology, YX; Software, YX; Validation, YZ, and YX.; Formal Analysis, JC and XC; Resources, MH; Data Curation, QL; Writing – Original Draft Preparation, QL; Writing – Review & Editing, MH; Visualization, YX; Supervision, YZ and QL; Project Administration, QL;

**Funding**: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest: Not Applicable.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Acknowledgements: Thank my colleagues in general ICU for cooperating with us in our prospective

study. Thanks to two public databases namely MIMIC and AmsterdamUMC for providing important data. Thanks for all patient advisers.

**Supplemental Figure:**

File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.

File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.

File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

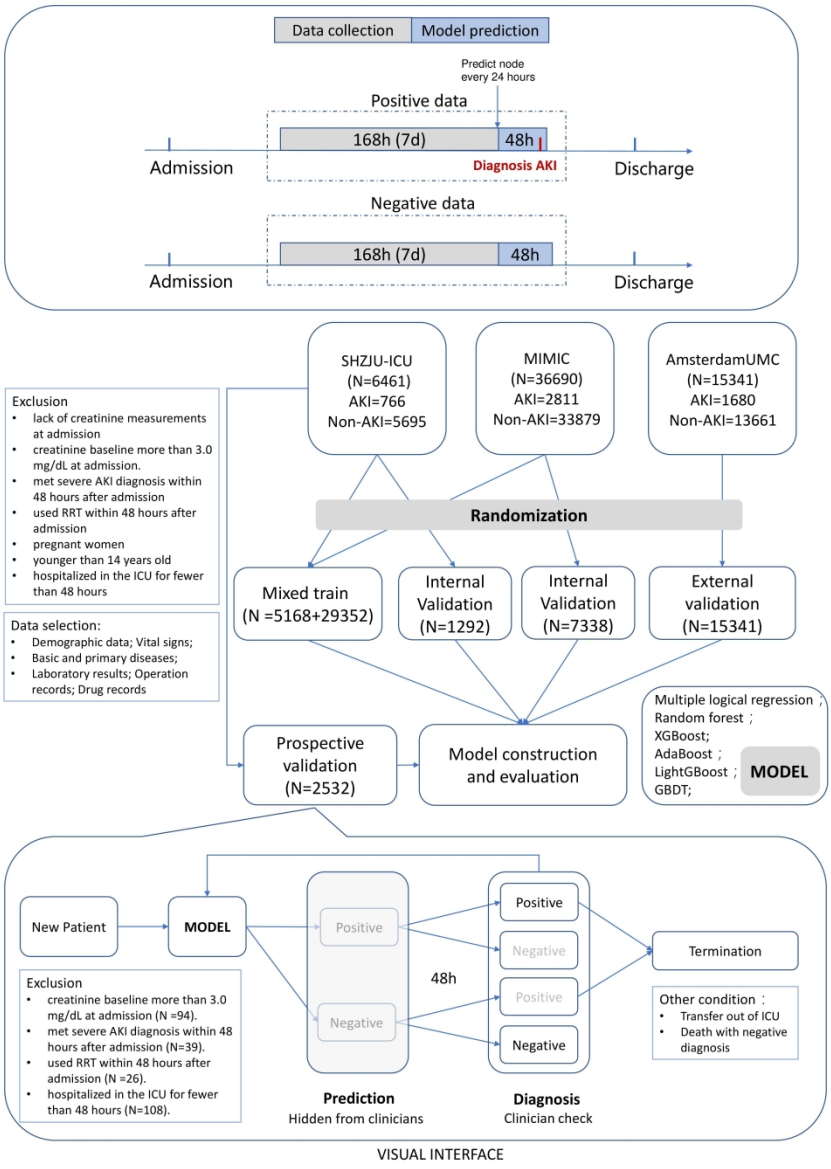
**Figure Legend:**

Figure 1: The research flow chart. The data collection time interval of the study was 7 days before

diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

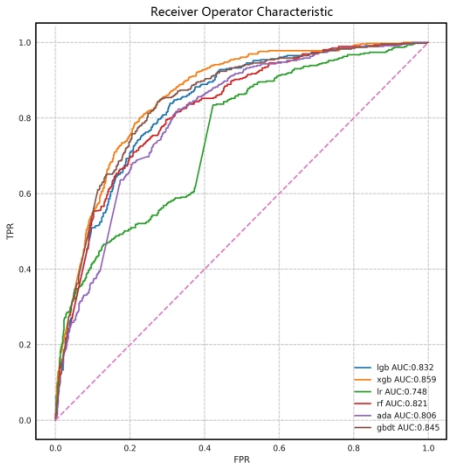
Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.



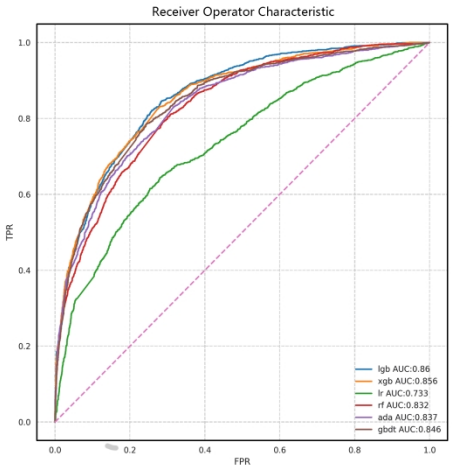

338x451mm (300 x 300 DPI)

Internal validation set ROC curves

A: SHZJU-ICU database validation set

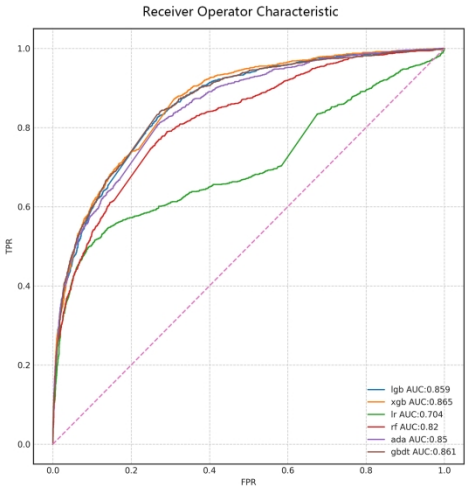



B: MIMIC database validation set

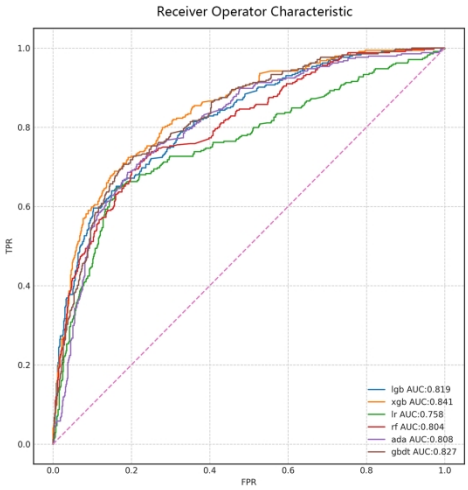


External and prospective validation set ROC curves

A: AmsterdamUMC database external validation set



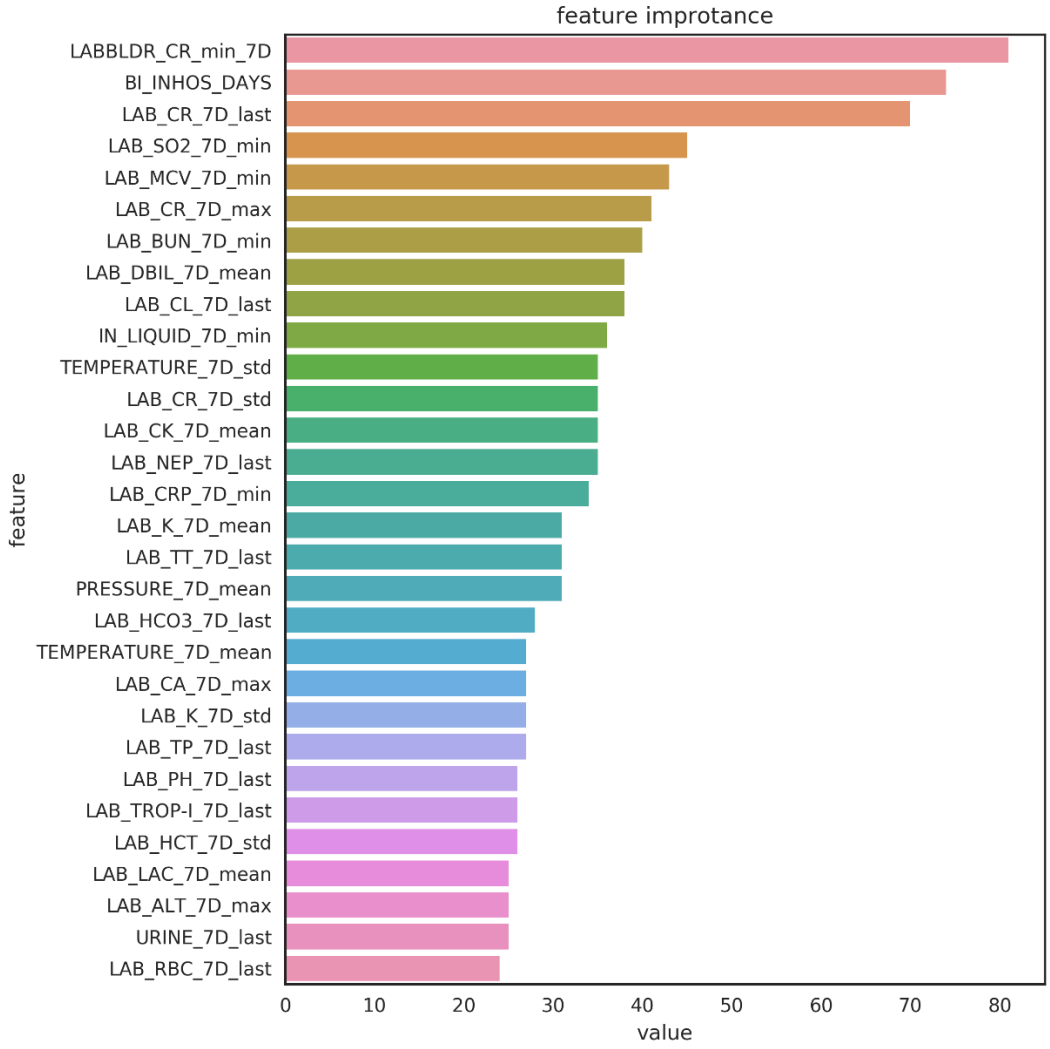

B: SHZJU-ICU database prospective validation set



**Supplemental Figure:**

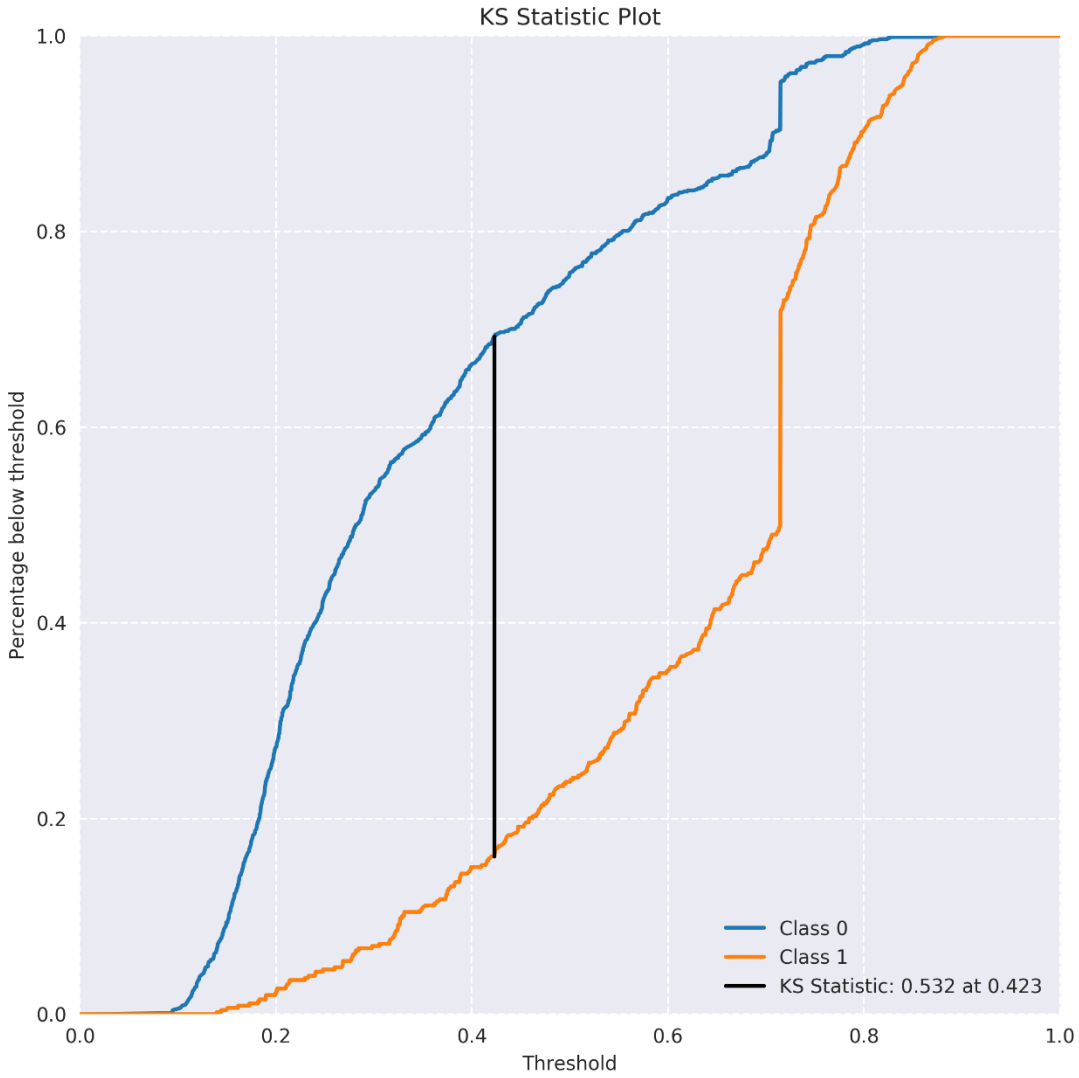
File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.




File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.



File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

| | SHZJU-ICU | MIMIC | AmsterdamUMC | Including in Model |
|---|---|---|---|---|
| **Age** | √ | √ | √ | √ |
| **Gender** | √ | √ | √ | √ |
| **Race** | √ | √ | × | × |
| **Primary disease** | √ | √ | × | × |
| **Comorbidity** | √ | √ | × | × |
| **Ventilation** | √ | √ | √ | √ |
| **Operation** | √ | √ | × | × |
| **ICU hours** | √ | √ | √ | √ |
| **Survived** | √ | √ | √ | √ |
| **Vital sign** | | | | |
| **Temperature** | √ | √ | √ | √ |
| **Systolic pressure** | √ | √ | √ | √ |
| **Diastolic pressure** | √ | √ | √ | √ |
| **Respiratory rate** | √ | √ | √ | √ |
| **Heart rate** | √ | √ | √ | √ |
| **Oxygen saturation** | √ | √ | √ | √ |
| **Urine** | √ | √ | √ | √ |
| **GCS score** | √ | × | × | × |
| **Laboratory results** | | | | |
| **Scr** | √ | √ | √ | √ |
| **BUN** | √ | √ | √ | √ |
| **RBC** | √ | √ | √ | √ |
| **MCV** | √ | √ | √ | √ |
| **Hb** | √ | √ | √ | √ |
| **WBC** | √ | √ | √ | √ |
| **NEUT** | √ | √ | √ | √ |
| **PLT** | √ | √ | √ | √ |
| **HCV** | √ | √ | × | × |
| **TBLB** | √ | √ | √ | √ |
| **DBLB** | √ | √ | √ | √ |
| **IBLB** | √ | √ | √ | √ |
| **CRP** | √ | √ | √ | √ |
| **PCT** | √ | × | × | × |
| **Serum kalium** | √ | √ | √ | √ |
| **Serum natrium** | √ | √ | √ | √ |
| **Serum chlorine** | √ | √ | √ | √ |

| | | | | |
|---|---|---|---|---|
| **PT** | √ | √ | √ | √ |
| **APTT** | √ | √ | √ | √ |
| **INR** | √ | √ | √ | √ |
| **CK** | √ | √ | √ | √ |
| **LDH** | √ | √ | √ | √ |
| **Troponin** | √ | √ | √ | √ |
| **Blood glucose** | √ | √ | √ | √ |
| **PH** | √ | √ | √ | √ |
| **Lactic acid** | √ | √ | √ | √ |
| **Anion gap** | √ | √ | √ | √ |
| **ABC** | √ | √ | √ | √ |
| **SBC** | √ | √ | √ | √ |
| **$PaCO_2$** | √ | √ | √ | √ |
| **$PaO_2$** | √ | √ | √ | √ |

SHZJU-ICU: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; MIMIC: Medical Information Mart for Intensive Care; APTT: Activated partial thromboplastin time; INR: International normalized ratio; LDH: Lactate dehydrogenase; $PaCO_2$: Partial pressure of carbon dioxide; $PaO_2$ Partial pressure of oxygen; MCV: Mean corpuscular volume; PH: Potential of hydrogen; WBC: White blood cell count; NEUT: Proportion of neutrophils; CK: Creatine kinase; Scr: Serum Creatinine; BUN: Blood urea nitrogen; RBC: Red blood cell count; Hb: Hemoglobin; PT: Prothrombin time; TBLB: Total bilirubin; DBLB: Direct bilirubin; IBLB: Indirect bilirubin; CRP: C-reactive protein; HCV: Hematocrit value; PCT: Procalcitonin; ABC: Actual Bicarbonate Radical; SBC: Standard Bicarbonate Radical; PLT: Platelet count

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | SHZJU-ICU Prospective Negative (N=2188) | SHZJU-ICU Prospective Positive (N=344) | P-value |
|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 61.3 [47.5-73.4] | 62.1 [51.2-70.5] | 0.03 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 1365 (62.4) | 207 (60.2) | 0.438 |
| **Race** | | | | | | |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 2 (0.001) | 0 | 1 |
| **Black** | 1 (0.02) | 0 | 1 | 0 | 0 | 1 |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 2185 (99.9) | 344(100) | 1 |
| **Other** | 0 | 1 (0.1) | 1 | 1 (0.001) | 0 | 1 |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 0.8 (0.32) | 0.9 (0.35) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (11) | 20 (10) | 0.007 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 230 (10.5) | 26 (7.6) | 0.102 |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 15 (0.7) | 3 (0.8) | 0.726 |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 98 (4.5) | 19 (5.7) | 0.403 |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 142 (6.5) | 29 (8.6) | 0.203 |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 324(14.8) | 56 (16.4) | 0.466 |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 1102 (50.4) | 239 (69.5) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 1234(56.4) | 160 (46.7) | 0.001 |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 58.7 [34.2-160.4] | 145.6 [68.5-314.8] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 1868 (85.4) | 178 (51.9) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

| 1Section/Topic | | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L80 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L84-86 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P5, L104 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L116 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L119 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L126 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P6, L159 |

| | | | | |
|---|---|---|---|---|
| | 0e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | NA |
| Risk groups | 11 | );V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 3a | );V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 3b | );V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 |
| | 3c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, Table2 |
| Model development | 4a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 4b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes |
| Model specification | 5a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 5b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | );V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | );V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L236 |
| Interpretation | 9a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L217 |
| | 9b | );V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L198 |
| Implications | 20 | );V | Discuss the potential clinical use of the model and implications for future research. | Yes, P9, L193 |
| **Other information** | | | | |
| Supplementary information | 21 | );V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | );V | Give the source of funding and the role of the funders for the present study. | NA |

TRIPOD Checklist: Prediction Model Development and Validation

| 1Section/Topic | Item | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L83-92 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L83-100 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P4, L103 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L125 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L130-138 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L139 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P7, L159 |
| | 10e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | Yes; P7, L137 |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 13a | D;V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 13b | D;V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 and SupTable2 |
| | 13c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, SupTable3 |
| Model development | 14a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 14b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes, SupTable1 |
| Model specification | 15a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 15b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | D;V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | D;V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L246 |
| Interpretation | 19a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L237 |
| | 19b | D;V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L210-245 |
| Implications | 20 | D;V | Discuss the potential clinical use of the model and implications for future research. | Yes, P12, L254 |
| **Other information** | | | | |
| Supplementary information | 21 | D;V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | D;V | Give the source of funding and the role of the funders for the present study. | NA |

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center prospective study



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2021-054092.R3 |
| Article Type: | Original research |
| Date Submitted by the Author: | 11-Dec-2021 |
| Complete List of Authors: | Liang, Qiqiang; Zhejiang University School of Medicine Second Affiliated Hospital<br>Xu, Yongfeng; Zhejiang University School of Medicine Second Affiliated Hospital<br>Zhou, Yu; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Xinyi; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Juan; Zhejiang University School of Medicine Second Affiliated Hospital<br>Huang, Man; Zhejiang University School of Medicine Second Affiliated Hospital, |
| Primary Subject Heading: | Intensive care |
| Secondary Subject Heading: | Diagnostics |
| Keywords: | Acute renal failure < NEPHROLOGY, INTENSIVE & CRITICAL CARE, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center prospective study

Qiqiang Liang[1]; Yongfeng Xu[1]; Yu Zhou[2]; Xinyi Chen[2], MM; Juan Chen[2]; Man Huang[1*]

General Intensive Care Unit, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

*Corresponding author: Man Huang, MD PhD.

Department: General Intensive Care Unit

Hospital: Second Affiliated Hospital of Zhejiang University School of Medicine

Address: No. 1511, Jianghong Road, Bingjiang District, Hangzhou City, Zhejiang Province

Tel: + 86 571 89713427

Fax: + 86 571 89713427

E-mail: huangman@zju.edu.cn.

ORCID: 0000-0002-0842-2254

**Abstract:**

Objectives: There are many studies of acute kidney injury (AKI) diagnosis models lack of external validation and prospective validation. We constructed the models using three databases to predict severe AKI within 48 hours in intensive care unit (ICU) patients.

Design: A retrospective and prospective cohort study.

Setting: We studied critically ill patients in our database (SHZJU-ICU) and two other public databases, the Medical Information Mart for Intensive Care (MIMIC) and AmsterdamUMC databases, including basic demographics, vital signs, and laboratory results. We predicted the diagnosis of severe AKI in patients in the next 48 hours using machine-learning algorithms with the three databases. Then, we carried out real-time severe AKI prediction in the prospective validation study at our center for one year.

Participants: All patients included in three databases with uniform exclusion criteria.

Primary and secondary outcome measures: Effect evaluation index of prediction models.

Results: We included 58492 patients, and a total of 5257 (9.0%) patients met the definition of severe AKI. In the internal validation of the SHZJU-ICU and MIMIC databases, the best area under the receiver operating characteristic curve (AUROC) of the model was 0.86. The external validation results by AmsterdamUMC database were also satisfactory, with the best AUROC of 0.86. A total of 2532 patients were admitted to the center for prospective validation; 358 positive results were predicted, and 344 patients were diagnosed with severe AKI, with the best sensitivity of 0.72, the specificity of 0.80, and the AUROC of 0.84.

Conclusion: The prediction model of severe AKI exhibits promises as a clinical application based on dynamic vital signs and laboratory results of multi-center databases with prospective and external validation.

**Keywords**: Machine learning; Acute kidney injury; Real-time prospective validation; External validation.

**Strengths and limitations of this study**:

A prospective validation in machine learning of AKI research rather than other studies;

Three large database containing different national populations and regions;

Variable's sampling limited by the monitoring frequency of clinical data;

Differences in the samples proportion of three databases;

The dimensions of variables are not rich enough.

**Data sharing statement:** No data available.

## Introduction:

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalization and mortality[1]. AKI is divided into three types according to the various aetiologies: prerenal (renal hypoperfusion), intrarenal (vascular, glomerular, or tubulointerstitial lesions), and postrenal (urinary tract obstruction)[2]. Although nearly all diseases associated with ICU admission may cause AKI, acute tubular necrosis (ATN) and prerenal azotaemia are the most common causes[3].

All AKI diagnostic criteria including the latest Kidney Disease: Improving Global Outcomes (KDIGO) standard are currently based on the creatinine level and urine volume[4]. However, the increase in the creatinine level or decrease in the urine volume lags the onset of AKI[2]. Many studies have suggested that early diagnosis and treatment of reversible AKI can reduce mortality[5]. Therefore, the creatinine level and urine volume are not satisfactory to meet clinical diagnostic demands. Consequently, many

researchers have tried to develop an early warning model by analysing the risk factors for AKI[6].

Patient complications, such as diabetes, hypertension, cardiovascular disease, chronic liver disease, sepsis, and trauma, are identified as important risk factors for AKI[7]. The AKI prediction model and scoring system developed based on high-risk factors has gradually become the focus of research considering the lower clinical application threshold compared with that of new biomarkers[6]. Although most previous prediction models use the multiple logistic regression model, a variety of AKI prediction models based on machine learning have resulted in satisfactory outcomes[6]. Since the first AKI prediction model study based on artificial intelligence was published in 2016, researchers have built more than 20 published AKI prediction models successively by using local or multi-centre databases[6,8-14]. The results indicate that these models can predict the occurrence of AKI and the need for renal replacement therapy (RRT) within 24 or 48 hours, with accuracies ranging from 81% to 97%[6,15]. In addition, many studies have focused on subspecialized conditions, including cardiac surgery, trauma, and burns[14-16]. However, the common defect in these studies is the lack of external validation and prospective validation, which causes the prediction model to deviate from the clinical scenarios and limits extrapolation beyond the scope of the data.

In this study, we built models to predict AKI within 48 hours in critically ill patients by using transcontinental three databases. Then, we evaluated the clinical effect of the model through a one-year prospective validation at our centre.

## Methods:

### Study design and setting

We collected patients using three ICU databases and prospectively validated the models in our centre.

The first database was our centre general ICU database (SHZJU-ICU) of the Second Affiliated Hospital of Zhejiang University School of Medicine, an academic teaching hospital. Since its establishment in 2017, it has included 12000 ICU patients' data and is updated daily. The Medical Information Mart for Intensive Care (MIMIC) III database, the second one, is an open ICU database provided by the Massachusetts Institute of Technology and includes nearly 60000 ICU patients from North America[17]. Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23000 patients admitted to ICUs in parts of Europe[18]. The research flow chart is shown in Figure 1.

**Study definition**

In this study, the diagnosis of AKI was confirmed based on three stages according to the KDIGO criteria[4]. We defined the patients who met the KDIGO AKI II and III criteria as severe AKI groups and the others as negative groups. We excluded patients with lack of creatinine measurements during admission, patients with creatinine baseline more than 3.0 mg/dL at admission, patients who met severe AKI diagnosis within 24 hours, and patients who used RRT within 48 hours after admission[19]. In addition, we excluded pregnant women, patients younger than 14 years old, and patients hospitalized in the ICU for fewer than 48 hours. After the patient was admitted to ICU, we performed a prediction every 24 hours and recorded a prediction time. If the patient was diagnosed with severe AKI within 48 hours, the predictive time was defined as a positive predictive point, and the others were defined as a negative point. The study was evaluated and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine as study number 2019-078.

**Data Collection**

The variables included demographic data, vital signs, basic and primary diseases, laboratory results,

important operation records, and drug records. Comorbidity included hypertension, diabetes, cardiopathy, liver disease, and malignant tumors. The primary disease was the main cause of admission to the ICU following the ICD-10 codes. The vital signs and clinical laboratory results were transformed into different variables according to the average, variance, maximum, minimum, and final value before diagnosis. We use a method similar to the forward incremental method in the multivariate logic regression model, that is, the combination of embedded feature selection and forward addition for feature selection. First of all, all variables are trained in the model, then list by variables importance. variables are added to the model one by one according to the variable importance. a variable is retained if it causes the AUC growth to be greater than 0.01, otherwise delete it. We transformed the MIMIC and AmsterdamUMC databases according to our centre database structure, unifying the unit and diagnostic codes. We deleted variables missing more than 50%. Variables missing more than 30% but less than 50% are listed to clinicians who determine the potential correlation between these variables and AKI. We carry out multiple interpolation for these variables which clinicians require to be retained, and the others deleted. Variables missing less than 30% are fill in multiple interpolation. All Missing data between three databases and values included in the model shown in supplementation file Table S1.

**Model construction and external validation**

The ratio of the training and internal validation sets was 4:1. The SHZJU and MIMIC databases training sets were mixed into a new training set. There were more negative data than positive data, so we randomly sampled the negative datasets and constructed a new data subset with a sampling ratio of positive and negative data of 1:5 in model building in order to extract the importance variables. In the subsequent model validation, we adopted the original data set. We used multiple logistic regression, random forest, XGBoost, AdaBoost, LightGBoost, gradient boosting decision tree (GBDT), and debug

to assess the variables and model-related parameters by the fivefold cross-validation method. After the models were built, we used the SHZJU-ICU and MIMIC test sets for internal validation of the model and the AmsterdamUMC database for external validation. The most appropriate cut-off value was determined according to the K-S curve. The prediction model represents the results of each prediction with a probability between 0 to 1.0. We define results more than 0.4 as high-risk, that is, positive results, and the rest as negative results. Through internal validation and external validation, we calibrated the model by adjusting the super-parameters and using the Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. All model building and validation processes were performed in Python 3.6.

**Prospective validation**

The prospective research period was 2020.01.01 to 2020.12.31. We collected real-time data when patients were admitted to the ICU, transformed the data according to the requirements, and formed a complete sample for the prediction model after passing the integrity test. We had established a visualization scheme and allowed researchers to review the predictions daily. The daily prediction results were not publicly accessible during the study to avoid affecting clinicians' decisions, but the diagnosis results were available to the researchers as visual graphics. We sampled the 20% predicted data every month and deleted samples with more than 50% missing values to ensure data correctness. When a patient has the following conditions, AKI prediction system will end the patient's prospective prediction: A. a positive diagnosis; B. Transfer out of ICU or death with negative diagnosis. All diagnosis of severe AKI needs to be reviewed by two ICU attending physicians independently, and if the they have different opinion, the third one will be appealed.

**Statistical analysis:**

The population characteristics were reported as the medians and inter-quartile ranges (IQRs) for skewed data and the means and standard deviations for normally distributed data. The independent sample T-test was used for normally distributed data, and the rank-sum test was used for the rest. Dichotomous variables were assessed by the chi-square test, and a P value less than 0.05 was considered statistically significant. The non-normally distributed data were analyzed by exponential transformation and logarithmic transformation. The effect of the model was evaluated by parameters such as the AUROC, accuracy, specificity, and F1-score.

**Patient and Public Involvement:**

The information of cases in three databases was in a state of complete desensitization in the process of building the model. During the prospective study, all the patients signed an informed consent form at the beginning of admission to ICU. The real-time data discussed and used by only the study members, and were not made public during the study period. All data were anonymized before the authors accessed them for the purpose of this study. Therefore, patients' priorities, experience, and preferences will not affect the development of the research question and outcome measures. If necessary, we will inform patients of relevant research results by telephone.

**Results:**

According to the inclusion criteria and exclusion criteria, we selected 58492 patients from three databases who met the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36690 patients from the MIMIC database, and 15341 patients from the AmsterdamUMC database. A total of 5257 (9.0%) patients met the definition of severe AKI (11.8% in SHZJU-ICU, 7.6% in MIMIC, and 10.9% in AmsterdamUMC). The distributions of age and sex in the three centers were similar, but

the differences in race were large. Asian patients accounted for more than 99% in the SHZJU-ICU database, and only approximately 2.5% in the MIMIC database. White people accounted for more than 70% of the MIMIC database. In addition, patients from the MIMIC database had a higher incidence of the tumor, liver cirrhosis, diabetes, and hypertension. Patients in the AmsterdamUMC and SHZJU-ICU databases had a higher proportion of mechanical ventilation and overall survival rate. Severe AKI patients had longer ICU hospital stays and higher mortality. More details are presented in Table 1.

Table 1: Clinical Demographics and Outcomes in patients with or without severe acute kidney injury.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | MIMIC Negative (N=33879) | MIMIC Severe AKI (N=2811) | P-value | AmsterdamUMC Negative (N=13661) | AmsterdamUMC severe AKI (N=1680) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 62.0 [45.0-76.0] | 64.0 [52.0-76.0] | <0.001 | 64.5 [54.5,74.5] | 64.3 [52.5,78.5] | 0.04 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 18976 (56.0) | 1524 (54.2) | 0.06 | 8920 (65.3) | 1054 (63.9) | 0.03 |
| **Race** | | | | | | | / | / | / |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 23796 (70.2) | 2005 (71.3) | 0.23 | / | / | / |
| **Black** | 1 (0.02) | 0 | 1 | 2455 (7.2) | 199 (7.1) | 0.45 | / | / | / |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 918 (2.7) | 70 (2.5) | 0.54 | / | / | / |
| **Other** | 0 | 1 (0.1) | 1 | 6710 (19.8) | 537 (19.1) | 0.34 | / | / | / |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 1.1 (0.41) | 1.1 (0.50) | 0.04 | 1.2 (0.35) | 1.1 (0.38) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (12) | 23 (12) | <0.001 | 18 (13) | 23 (10) | <0.001 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 6652 (19.6) | 752 (26.8) | <0.001 | / | / | / |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 1180 (3.5) | 426 (15.2) | <0.001 | / | / | / |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 4840 (14.3) | 336 (12.0) | 0.001 | / | / | / |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 7250 (21.4) | 789 (28.1) | <0.001 | / | / | / |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 16328 (48.2) | 1513 (53.8) | <0.001 | / | / | / |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 13009 (38.4) | 1503(53.5) | <0.001 | 7718(56.5) | 1248(74.3) | <0.001 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 12455 (36.7) | 1430 (50.8) | <0.001 | / | / | / |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 51.0 [28.5-110.3] | 103.4 [50.1-261.3] | <0.001 | 24.0 [19.8,63.5] | 142.0 [45.4,394.2] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 25648 (75.7) | 1356 (48.2) | <0.001 | 10942(80.1) | 1033 (61.5) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

There were significant differences in the important parameters of the variables among the different models (see Figure S1). However, the trend of the creatinine level in the past week was still an important variable, followed by urine volume, blood urea nitrogen level, temperature, and length of ICU stay. The cut-off value used to distinguish between a negative and positive prediction was determined by the K-S curve, with value of 0.423 (see Figure S2). The GBDT model had the best prediction effect in the test set, followed by XGBoost and LightGBoost. In the two central internal validation sets, the two best-performing machine learning algorithms with great AUROC are LightGBoost (SHZJU-ICU of 83.2%, MIMIC of 86.0%) and XGBoost (SHZJU-ICU 85.9%, MIMIC 85.6%), as detailed in Figure 2. Overall, the sensitivity (SHZJU-ICU 0.84, MIMIC 0.83) and the negative predictive value (SHZJU-ICU 0.90, MIMIC 0.90) of the predictive model were high, but the specificity was general (SHZJU-ICU 0.79, MIMIC 0.75), as shown in Table 2. In the external validation based on AmsterdamUMC database, the overall model validation effect results were satisfactory, and XGBoost had the best performance, with an AUROC of 0.84, as shown in Figure 2 and Table 2.

Table 2: Model validation results by three databases with machine learning algorithm

| *Model* | *AUROC* | *Accuracy* | *Sensitivity* | *Specificity* | *PPV* | *NPV* | *F1* |
|---|---|---|---|---|---|---|---|
| ***Internal validation with SHZJU-ICU database*** | | | | | | | |
| ***Logistic regression*** | *0.748* | *0.662* | *0.834* | *0.576* | *0.496* | *0.874* | *0.622* |
| ***LightGBoost*** | *0.832* | *0.741* | *0.839* | *0.692* | *0.576* | *0.896* | *0.683* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ***GBDT*** | *0.845* | *0.765* | *0.843* | *0.725* | *0.606* | *0.902* | *0.705* |
| ***AdaBoost*** | *0.806* | *0.721* | *0.824* | *0.67* | *0.555* | *0.884* | *0.663* |
| ***Random Forest*** | *0.821* | *0.763* | *0.71* | *0.789* | *0.627* | *0.845* | *0.666* |
| ***XGBoost*** | ***0.859*** | *0.779* | *0.81* | *0.763* | *0.631* | *0.889* | *0.709* |
| ***Internal validation with MIMIC database*** | | | | | | | |
| ***Logistic regression*** | *0.733* | *0.695* | *0.643* | *0.72* | *0.535* | *0.801* | *0.584* |
| ***LightGBoost*** | ***0.86*** | *0.768* | *0.822* | *0.741* | *0.613* | *0.893* | *0.702* |
| ***GBDT*** | *0.846* | *0.765* | *0.786* | *0.755* | *0.616* | *0.876* | *0.691* |
| ***AdaBoost*** | *0.837* | *0.732* | *0.831* | *0.683* | *0.567* | *0.89* | *0.674* |
| ***Random Forest*** | *0.832* | *0.738* | *0.791* | *0.712* | *0.578* | *0.872* | *0.668* |
| ***XGBoost*** | *0.856* | *0.758* | *0.833* | *0.721* | *0.598* | *0.895* | *0.695* |
| ***External validation with AmsterdamUMC database*** | | | | | | | |
| ***Logistic regression*** | *0.704* | *0.767* | *0.516* | *0.893* | *0.706* | *0.787* | *0.596* |
| ***LightGBoost*** | *0.859* | *0.763* | *0.827* | *0.731* | *0.606* | *0.894* | *0.7* |
| ***GBDT*** | *0.861* | *0.764* | *0.84* | *0.727* | *0.606* | *0.901* | *0.704* |
| ***AdaBoost*** | *0.85* | *0.755* | *0.813* | *0.726* | *0.597* | *0.886* | *0.689* |
| ***Random Forest*** | *0.82* | *0.743* | *0.77* | *0.729* | *0.587* | *0.864* | *0.666* |
| ***XGBoost*** | ***0.865*** | *0.75* | *0.873* | *0.688* | *0.584* | *0.916* | *0.7* |
| ***Prospective validation with SHZJU-ICU*** | | | | | | | |
| ***Logistic regression*** | *0.758* | *0.772* | *0.648* | *0.834* | *0.662* | *0.826* | *0.655* |
| ***LightGBoost*** | *0.819* | *0.796* | *0.596* | *0.895* | *0.74* | *0.816* | *0.66* |
| ***GBDT*** | *0.827* | *0.781* | *0.706* | *0.818* | *0.66* | *0.848* | *0.683* |
| ***AdaBoost*** | *0.808* | *0.766* | *0.686* | *0.805* | *0.638* | *0.837* | *0.661* |
| ***Random Forest*** | *0.804* | *0.755* | *0.715* | *0.775* | *0.613* | *0.845* | *0.66* |
| ***XGBoost*** | ***0.841*** | *0.779* | *0.724* | *0.807* | *0.652* | *0.854* | *0.686* |

*AUROC*: Area Under Receiver Operating Characteristic; *PPV*: Positive Predictive Value; *NPV*: Negative Predictive Value; *SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *GBDT*: Gradient Boosted Decision Tree.

According to the inclusion and exclusion criteria, we delete 267 patients among 94 patients with creatinine baseline more than 3.0 mg/dL at admission, 39 patients met severe AKI diagnosis within 24 hours, and 26 patients who used RRT within 48 hours after admission, 108 patients hospitalized in the ICU for fewer than 48 hours. A total of 2532 patients were admitted to our centre for prospective validation, and the prediction model made 16858 times predictions. In the prospective cohort, there was no significant difference in age, gender, baseline creatinine and urea nitrogen, and complications. The

proportion of mechanical ventilation and the ICU stay time in AKI patients were longer with higher mortality. Above all, there was no significant difference between the prospective and the retrospective cohort. More detail sees in supplementation file Table S2. In the end, 358 positive results were predicted, and the rest were negative results. There are 344 patients with severe AKI were diagnosed and the prediction accuracy was 83.5%. The model with the highest area under the curve was XGBoost, 0.84 with the best sensitivity of 0.72, the specificity of 0.80. The results of the prospective study are similar to those of the external validation of the model, and are relatively stable. More detail is presented in Figure 3 and Table 2.

**Discussion:**

In this study, we built predictive models by machine learning to predict the incidence of severe AKI with three databases in different regions and in the next 48 hours. After internal and external validation, prospective validation over one year was carried out to verify the model effects. The three databases come from three countries that are in Asia, Europe, and North America, which proves that the model is universal to some extent.

Despite the huge amount of data, many databases are still not suitable for prospective research because they are not updated promptly. Tomasev, N and colleagues have provided research on AKI prediction models with a large amount of data[10]. The study covered 703782 adult patients with 6 billion individual items, including 620000 elements. In this study, a depth neural network model was used for real-time prediction. A total of 55.8% of severe AKI patients were predicted within the first 48 hours, although each accurate prediction was accompanied by two mispredictions[10]. This study provided a new scheme for real-time prediction and indicated that we should prospectively evaluate and

independently validate models to explore their effectiveness. In a prospective study, Flechet, M et al compared an AKI prediction model with clinicians in 252 patients and found that the clinical effect of the random forest model for predicting AKI-II/III was equivalent to that of clinicians. Our prediction model graphical visualization of the model was installed in the centre's database for better usage. In addition, our database is updated daily to achieve daily predictions and present the results to researchers. In the prospective validation of our study, the stability of the prediction model confirmed its promise, which provides a basis for future research.

There are many studies of artificial intelligence for predicting the occurrence of AKI, but most of them are single-centre studies, and the extrapolation effect has been controversial. In 2016, Koyner, J, and his colleague published the first study of an AKI prediction model based on multi-centre data. In all 202961 patients, 17541 (8.6%) had AKI, 4251 (3.5%) had AKI-II, and 1242 (0.6%) had AKI-III. A multivariate logistic regression model was used to predict AKI in this study with an AUROC of 0.74 (95% CI=0.74). With the classification of AKI, the AUROC of the prediction model gradually increased to 0.84[8]. Subsequently, the study team used a new machine learning algorithm, to build a more accurate model to predict the occurrence of AKI-II, with an AUROC of 0.9 within 24 hours and 0.87 within 48 hours[19]. Recently, the research team included data from two other centres, namely, LUMC (N=200613) and NUS (N=246895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85 to 0.86, suggesting that the artificial intelligence model has stable predictive ability[12]. This series of studies included many data points, suggesting the feasibility of artificial intelligence in the diagnosis of AKI, but the proportion of positive patients (3.5%) and ICU patients (30%) was too low to properly predict AKI. Our research is similar to the above. The SAHZJU-ICU database is a single-centre database representing south-eastern China, and the MIMIC database is a well-known open ICU database in the

United States. The AmsterdamUMC database is a public database located in Europe. The population structure and diseases in the three databases are complete but different in the distribution of complications and race. Therefore, it provides a prediction model with unparalleled stability compared with other studies.

Limitations:

This retrospective multi-centre study was unable to carry out more clinical feature mining and comparison because of different data structures. The differences between the three databases partly reflect some demographic differences between Europe, the United States and China, resulting in a decline in the accuracy of the prediction model. There are some differences in the number of patients included in the three databases, which may affect the choice of variables. As a result of the study design, we deleted patients with ICU hospitalization of less than 48 hours, which may result in the exclusion of most relatively mild patients and may reduce false positives. Second, in the prospective data study in 2020, there may be deviations in the inclusion of patients in the centre, thus affecting the interpretation of the follow-up prospective results. Finally, given the low incidence of severe AKI and the great difference in the proportion of positive and negative samples, the data may be accidental. Our model seems to be superior to diagnostic non-AKI patients rather than AKI because of the proportion of positive data that we include. In the retrospective study, we reduced the proportion of negative data by randomization but retain all date in prospective phase with the sensitivity decreases.

**Conclusion:**

Based on databases of patients of different races from different countries, we constructed stable

machine learning models to predict the occurrence of AKI in the next 48 hours. Prospective validation through the implementation of an updated local database is an effective exploration of further research.

Data sharing statement: No data available.

Ethics approval: Our research was approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

Competing interests for authors: The authors declare that they have no competing interests.

**Reference:**

1. Ronco, C.; Bellomo, R.; Kellum, J.A. Acute kidney injury. *The Lancet* **2019**, *394*, 1949-1964, doi:10.1016/s0140-6736(19)32563-2.
2. Pakula, A.M.; Skinner, R.A. Acute Kidney Injury in the Critically Ill Patient: A Current Review of the Literature. *J Intensive Care Med* **2016**, *31*, 319-324, doi:10.1177/0885066615575699.
3. Bellomo, R.; Kellum, J.A.; Ronco, C. Acute kidney injury. *The Lancet* **2012**, *380*, 756-766, doi:10.1016/s0140-6736(11)61454-2.
4. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* **2012**, *120*, c179-184, doi:10.1159/000339789.
5. Coelho, S.; Cabral, G.; Lopes, J.A.; Jacinto, A. Renal regeneration after acute kidney injury. *Nephrology (Carlton)* **2018**, *23*, 805-814, doi:10.1111/nep.13256.
6. Wilson, F.P.; Greenberg, J.H. Acute Kidney Injury in Real Time: Prediction, Alerts, and Clinical Decision Support. *Nephron* **2018**, *140*, 116-119, doi:10.1159/000492064.

7. Wu, Y.; Hao, W.; Chen, Y.; Chen, S.; Liu, W.; Yu, F.; Hu, W.; Liang, X. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* **2020**, *42*, 1127-1134, doi:10.1080/0886022X.2020.1843491.

8. Koyner, J.L.; Adhikari, R.; Edelson, D.P.; Churpek, M.M. Development of a Multicenter Ward-Based AKI Prediction Model. *Clin J Am Soc Nephrol* **2016**, *11*, 1935-1943, doi:10.2215/CJN.00280116.

9. Low, S.; Vathsala, A.; Murali, T.M.; Pang, L.; MacLaren, G.; Ng, W.Y.; Haroon, S.; Mukhopadhyay, A.; Lim, S.L.; Tan, B.H., et al. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* **2019**, *20*, 32, doi:10.1186/s12882-019-1206-4.

10. Tomasev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I., et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116-119, doi:10.1038/s41586-019-1390-1.

11. Flechet, M.; Falini, S.; Bonetti, C.; Güiza, F.; Schetz, M.; Van den Berghe, G.; Meyfroidt, G. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Critical Care* **2019**, *23*, 282, doi:10.1186/s13054-019-2563-x.

12. Churpek, M.M.; Carey, K.A.; Edelson, D.P.; Singh, T.; Astor, B.C.; Gilbert, E.R.; Winslow, C.; Shah, N.; Afshar, M.; Koyner, J.L. Internal and External Validation of a Machine Learning Risk Score for Acute Kidney Injury. *JAMA Netw Open* **2020**, *3*, e2012892, doi:10.1001/jamanetworkopen.2020.12892.

13. Yuan, Q.; Zhang, H.; Deng, T.; Tang, S.; Yuan, X.; Tang, W.; Xie, Y.; Ge, H.; Wang, X.; Zhou, Q., et al. Role of Artificial Intelligence in Kidney Disease. *Int J Med Sci* **2020**, *17*, 970-984, doi:10.7150/ijms.42078.

14. Li, Y.; Xu, J.; Wang, Y.; Zhang, Y.; Jiang, W.; Shen, B.; Ding, X. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* **2020**, *43*, 752-761, doi:10.1002/clc.23377.

15. De Vlieger, G.; Kashani, K.; Meyfroidt, G. Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review. *Curr Opin Crit Care* **2020**, *26*, 563-573, doi:10.1097/MCC.0000000000000775.

16. James, M.T.; Pannu, N.; Hemmelgarn, B.R.; Austin, P.C.; Tan, Z.; McArthur, E.; Manns, B.J.; Tonelli, M.; Wald, R.; Quinn, R.R., et al. Derivation and External Validation of Prediction Models for Advanced Chronic Kidney Disease Following Acute Kidney Injury. *JAMA* **2017**, *318*, 1787-1797, doi:10.1001/jama.2017.16326.

17. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.W.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.A.-O.; Celi, L.A.; Mark, R.A.-O. MIMIC-III, a freely accessible critical care database.

18. Thoral, P.J.; Peppink, J.M.; Driessen, R.H.; Sijbrands, E.J.G.; Kompanje, E.J.O.; Kaplan, L.; Bailey, H.; Kesecioglu, J.; Cecconi, M.; Churpek, M., et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. *Crit Care Med*

**2021**, 10.1097/CCM.0000000000004916, doi:10.1097/CCM.0000000000004916.

19. Koyner, J.L.; Carey, K.A.; Edelson, D.P.; Churpek, M.M. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit Care Med* **2018**, *46*, 1070-1077, doi:10.1097/CCM.0000000000003123.

**Footnotes**

Contributions: Conceptualization, QL and MH; Methodology, YX; Software, YX; Validation, YZ, and YX.; Formal Analysis, JC and XC; Resources, MH; Data Curation, QL; Writing – Original Draft Preparation, QL; Writing – Review & Editing, MH; Visualization, YX; Supervision, YZ and QL; Project Administration, QL;

**Funding**: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest: Not Applicable.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Acknowledgements: Thank my colleagues in general ICU for cooperating with us in our prospective study. Thanks to two public databases namely MIMIC and AmsterdamUMC for providing important data. Thanks for all patient advisers.

**Supplemental Figure:**

File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.

File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.

File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

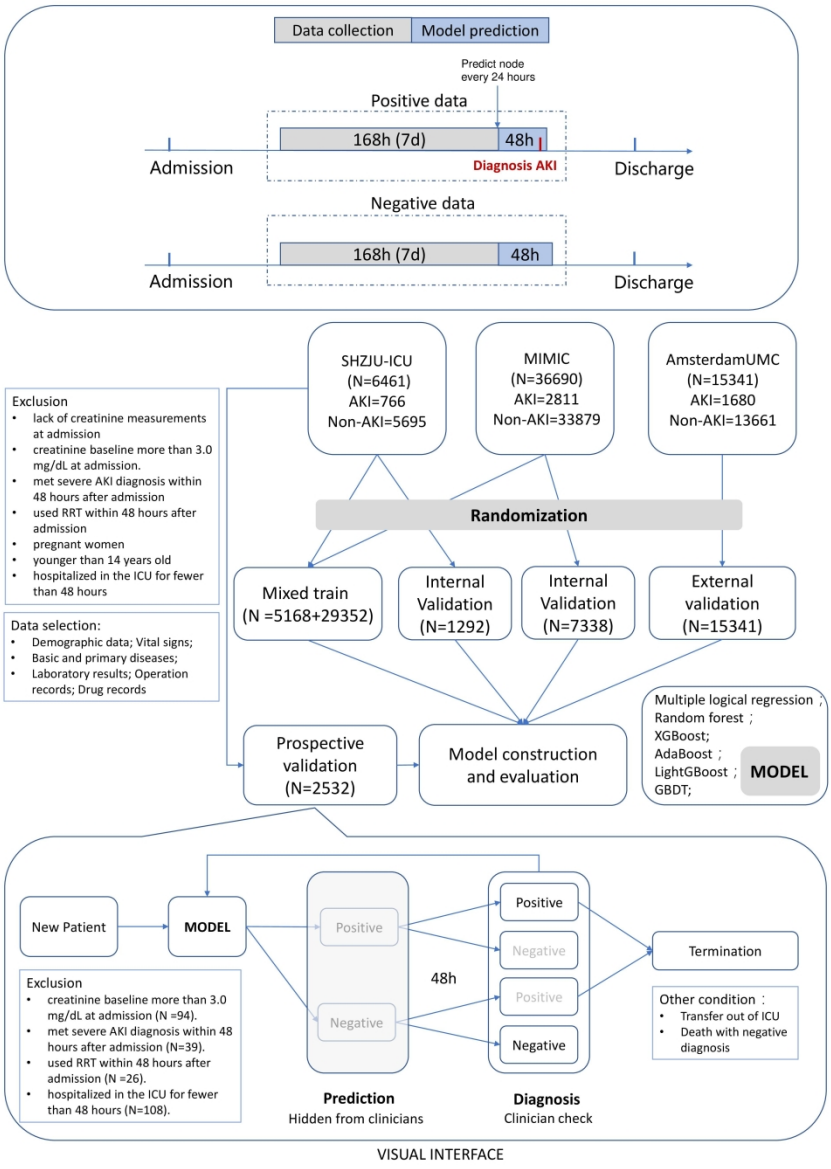
**Figure Legend:**

Figure 1: The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC

database.

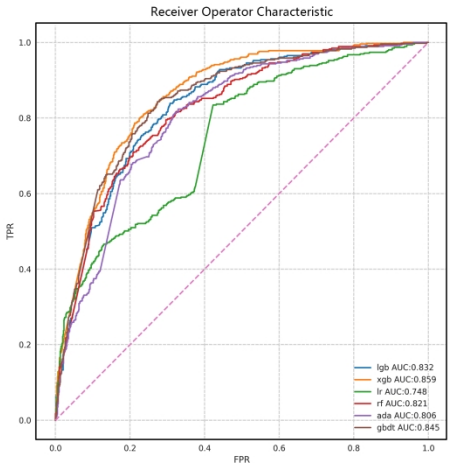
Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.



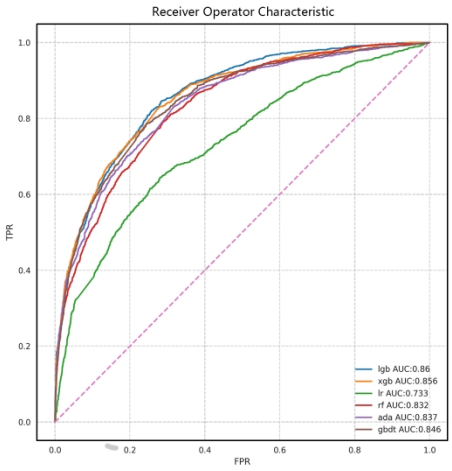

338x451mm (300 x 300 DPI)

Internal validation set ROC curves

A: SHZJU-ICU database validation set

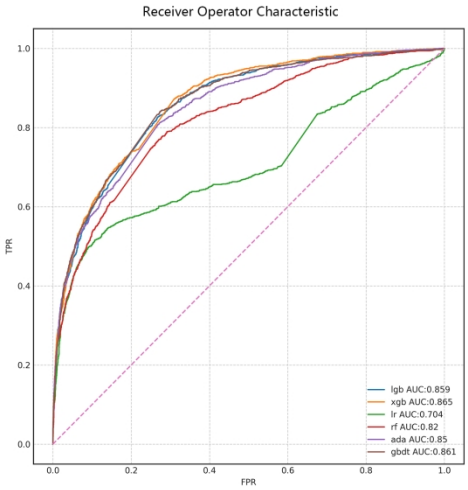



B: MIMIC database validation set

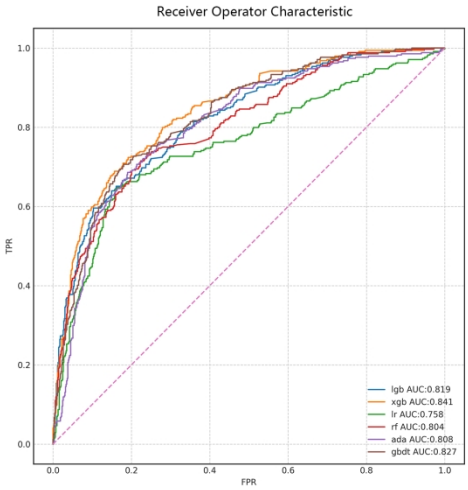


External and prospective validation set ROC curves

A: AmsterdamUMC database external validation set



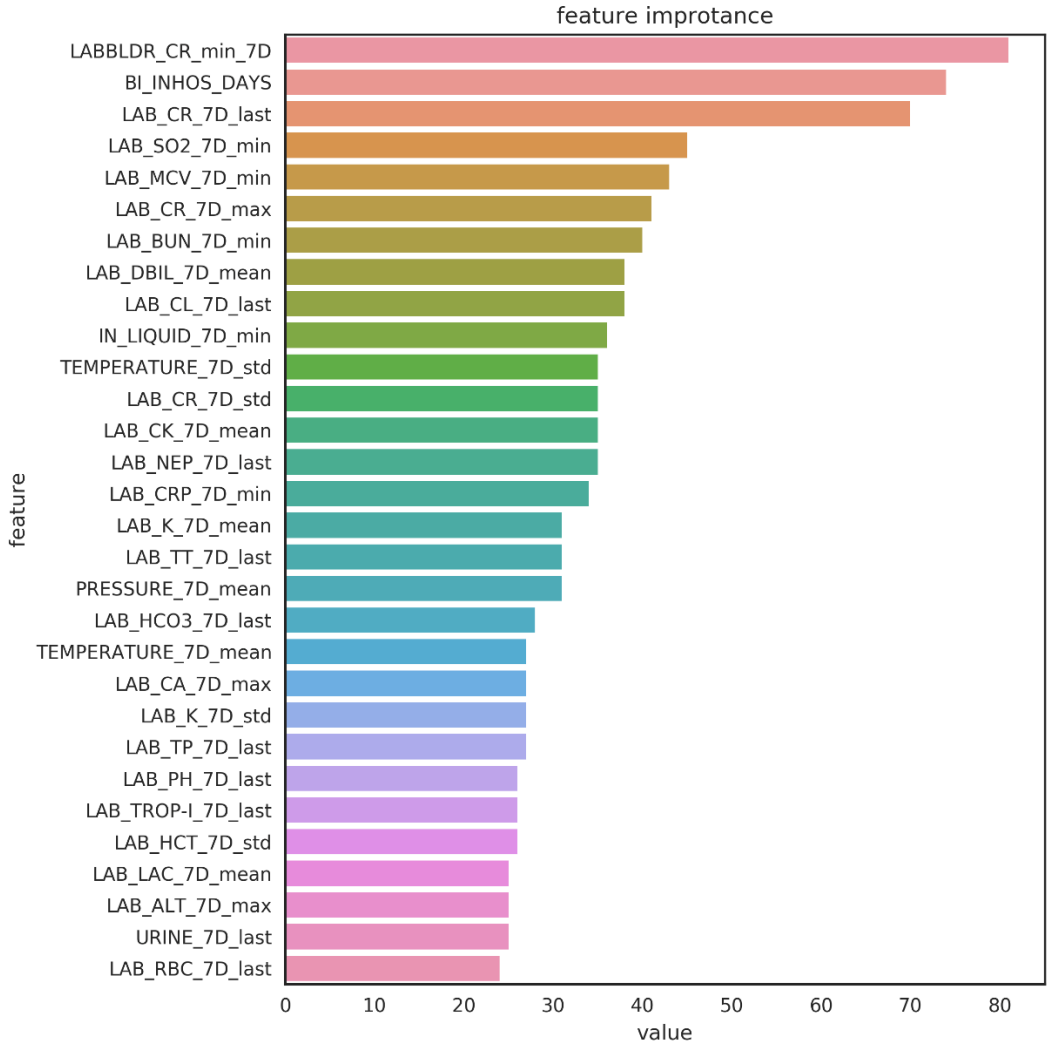

B: SHZJU-ICU database prospective validation set



**Supplemental Figure:**

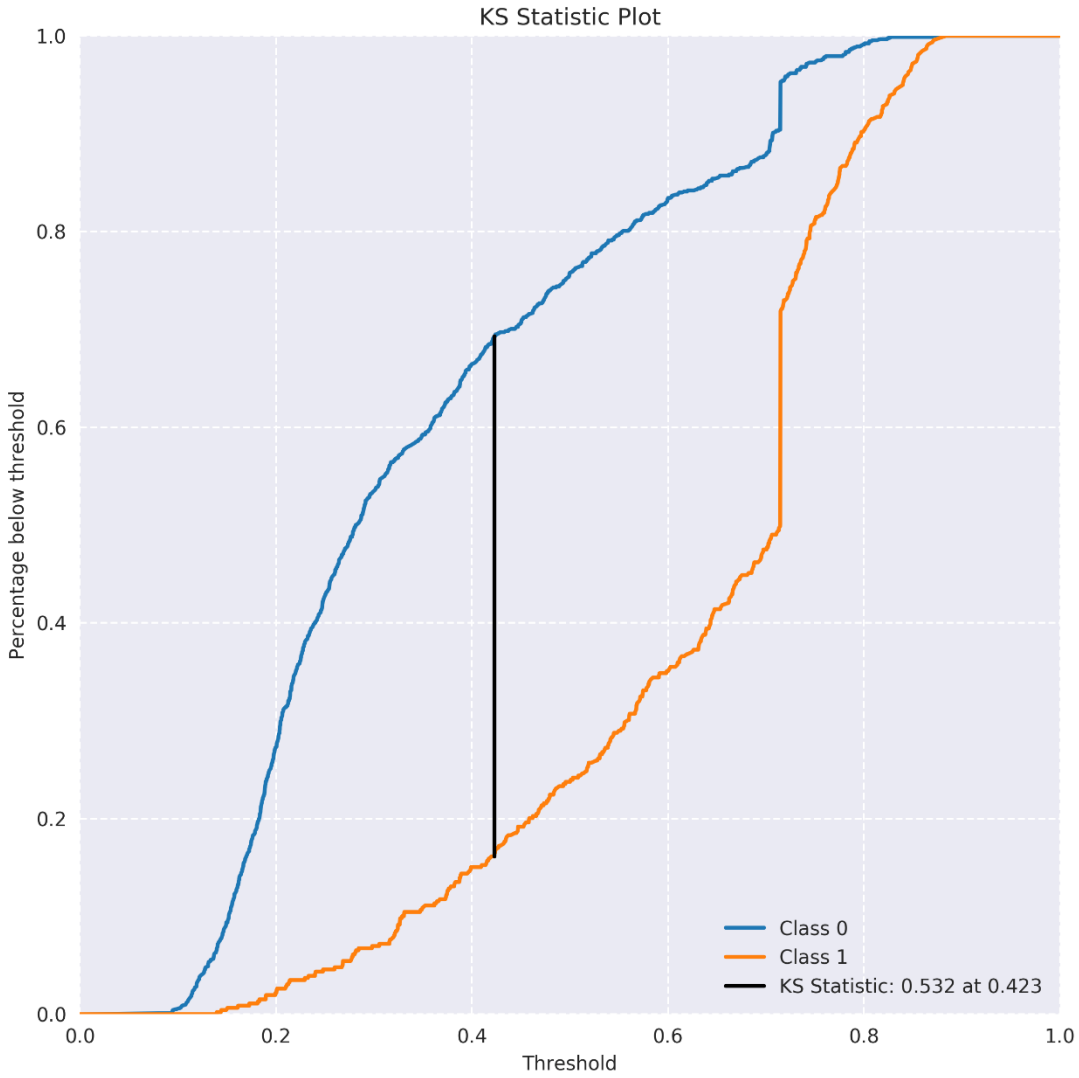
File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.




File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.



File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

| | SHZJU-ICU | MIMIC | AmsterdamUMC | Including in Model |
|---|---|---|---|---|
| **Age** | √ | √ | √ | √ |
| **Gender** | √ | √ | √ | √ |
| **Race** | √ | √ | × | × |
| **Primary disease** | √ | √ | × | × |
| **Comorbidity** | √ | √ | × | × |
| **Ventilation** | √ | √ | √ | √ |
| **Operation** | √ | √ | × | × |
| **ICU hours** | √ | √ | √ | √ |
| **Survived** | √ | √ | √ | √ |
| **Vital sign** | | | | |
| **Temperature** | √ | √ | √ | √ |
| **Systolic pressure** | √ | √ | √ | √ |
| **Diastolic pressure** | √ | √ | √ | √ |
| **Respiratory rate** | √ | √ | √ | √ |
| **Heart rate** | √ | √ | √ | √ |
| **Oxygen saturation** | √ | √ | √ | √ |
| **Urine** | √ | √ | √ | √ |
| **GCS score** | √ | × | × | × |
| **Laboratory results** | | | | |
| **Scr** | √ | √ | √ | √ |
| **BUN** | √ | √ | √ | √ |
| **RBC** | √ | √ | √ | √ |
| **MCV** | √ | √ | √ | √ |
| **Hb** | √ | √ | √ | √ |
| **WBC** | √ | √ | √ | √ |
| **NEUT** | √ | √ | √ | √ |
| **PLT** | √ | √ | √ | √ |
| **HCV** | √ | √ | × | × |
| **TBLB** | √ | √ | √ | √ |
| **DBLB** | √ | √ | √ | √ |
| **IBLB** | √ | √ | √ | √ |
| **CRP** | √ | √ | √ | √ |
| **PCT** | √ | × | × | × |
| **Serum kalium** | √ | √ | √ | √ |
| **Serum natrium** | √ | √ | √ | √ |
| **Serum chlorine** | √ | √ | √ | √ |

| | | | | |
|---|---|---|---|---|
| **PT** | √ | √ | √ | √ |
| **APTT** | √ | √ | √ | √ |
| **INR** | √ | √ | √ | √ |
| **CK** | √ | √ | √ | √ |
| **LDH** | √ | √ | √ | √ |
| **Troponin** | √ | √ | √ | √ |
| **Blood glucose** | √ | √ | √ | √ |
| **PH** | √ | √ | √ | √ |
| **Lactic acid** | √ | √ | √ | √ |
| **Anion gap** | √ | √ | √ | √ |
| **ABC** | √ | √ | √ | √ |
| **SBC** | √ | √ | √ | √ |
| **$PaCO_2$** | √ | √ | √ | √ |
| **$PaO_2$** | √ | √ | √ | √ |

SHZJU-ICU: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; MIMIC: Medical Information Mart for Intensive Care; APTT: Activated partial thromboplastin time; INR: International normalized ratio; LDH: Lactate dehydrogenase; $PaCO_2$: Partial pressure of carbon dioxide; $PaO_2$ Partial pressure of oxygen; MCV: Mean corpuscular volume; PH: Potential of hydrogen; WBC: White blood cell count; NEUT: Proportion of neutrophils; CK: Creatine kinase; Scr: Serum Creatinine; BUN: Blood urea nitrogen; RBC: Red blood cell count; Hb: Hemoglobin; PT: Prothrombin time; TBLB: Total bilirubin; DBLB: Direct bilirubin; IBLB: Indirect bilirubin; CRP: C-reactive protein; HCV: Hematocrit value; PCT: Procalcitonin; ABC: Actual Bicarbonate Radical; SBC: Standard Bicarbonate Radical; PLT: Platelet count

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | SHZJU-ICU Prospective Negative (N=2188) | SHZJU-ICU Prospective Positive (N=344) | P-value |
|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 61.3 [47.5-73.4] | 62.1 [51.2-70.5] | 0.03 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 1365 (62.4) | 207 (60.2) | 0.438 |
| **Race** | | | | | | |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 2 (0.001) | 0 | 1 |
| **Black** | 1 (0.02) | 0 | 1 | 0 | 0 | 1 |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 2185 (99.9) | 344(100) | 1 |
| **Other** | 0 | 1 (0.1) | 1 | 1 (0.001) | 0 | 1 |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 0.8 (0.32) | 0.9 (0.35) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (11) | 20 (10) | 0.007 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 230 (10.5) | 26 (7.6) | 0.102 |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 15 (0.7) | 3 (0.8) | 0.726 |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 98 (4.5) | 19 (5.7) | 0.403 |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 142 (6.5) | 29 (8.6) | 0.203 |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 324(14.8) | 56 (16.4) | 0.466 |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 1102 (50.4) | 239 (69.5) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 1234(56.4) | 160 (46.7) | 0.001 |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 58.7 [34.2-160.4] | 145.6 [68.5-314.8] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 1868 (85.4) | 178 (51.9) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

File name: **Table S3.docx**

Title of data: TRIPOD Checklist of Prediction Model Development and Validation

Description of data: Writing norms of this type of research.

| 1Section/Topic | | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L80 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L84-86 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P5, L104 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L116 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L119 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L126 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P6, L159 |

| | | | | |
|---|---|---|---|---|
| | 0e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | NA |
| Risk groups | 11 | );V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 3a | );V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 3b | );V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 |
| | 3c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, Table2 |
| Model development | 4a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 4b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes |
| Model specification | 5a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 5b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | );V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | );V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L236 |
| Interpretation | 9a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L217 |
| | 9b | );V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L198 |
| Implications | 20 | );V | Discuss the potential clinical use of the model and implications for future research. | Yes, P9, L193 |
| **Other information** | | | | |
| Supplementary information | 21 | );V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | );V | Give the source of funding and the role of the funders for the present study. | NA |

TRIPOD Checklist: Prediction Model Development and Validation

| 1Section/Topic | Item | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L83-92 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L83-100 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P4, L103 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L125 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L130-138 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L139 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P7, L159 |
| | 10e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | Yes; P7, L137 |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 13a | D;V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 13b | D;V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 and SupTable2 |
| | 13c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, SupTable3 |
| Model development | 14a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 14b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes, SupTable1 |
| Model specification | 15a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 15b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | D;V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | D;V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L246 |
| Interpretation | 19a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L237 |
| | 19b | D;V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L210-245 |
| Implications | 20 | D;V | Discuss the potential clinical use of the model and implications for future research. | Yes, P12, L254 |
| **Other information** | | | | |
| Supplementary information | 21 | D;V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | D;V | Give the source of funding and the role of the funders for the present study. | NA |

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center prospective study



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2021-054092.R4 |
| Article Type: | Original research |
| Date Submitted by the Author: | 04-Feb-2022 |
| Complete List of Authors: | Liang, Qiqiang; Zhejiang University School of Medicine Second Affiliated Hospital<br>Xu, Yongfeng; Zhejiang University School of Medicine Second Affiliated Hospital<br>Zhou, Yu; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Xinyi; Zhejiang University School of Medicine Second Affiliated Hospital<br>Chen, Juan; Zhejiang University School of Medicine Second Affiliated Hospital<br>Huang, Man; Zhejiang University School of Medicine Second Affiliated Hospital, |
| Primary Subject Heading: | Intensive care |
| Secondary Subject Heading: | Diagnostics |
| Keywords: | Acute renal failure < NEPHROLOGY, INTENSIVE & CRITICAL CARE, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Severe Acute Kidney Injury Predicting Model based on transcontinental databases: a single-center prospective study

Qiqiang Liang[1]; Yongfeng Xu[1]; Yu Zhou[2]; Xinyi Chen[2], MM; Juan Chen[2]; Man Huang[1*]

General Intensive Care Unit, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China.

*Corresponding author: Man Huang, MD PhD.

Department: General Intensive Care Unit

Hospital: Second Affiliated Hospital of Zhejiang University School of Medicine

Address: No. 1511, Jianghong Road, Bingjiang District, Hangzhou City, Zhejiang Province

Tel: + 86 571 89713427

Fax: + 86 571 89713427

E-mail: huangman@zju.edu.cn.

ORCID: 0000-0002-0842-2254

**Abstract:**

Objectives: There are many studies of acute kidney injury (AKI) diagnosis models lack of external validation and prospective validation. We constructed the models using three databases to predict severe AKI within 48 hours in intensive care unit (ICU) patients.

Design: A retrospective and prospective cohort study.

Setting: We studied critically ill patients in our database (SHZJU-ICU) and two other public databases, the Medical Information Mart for Intensive Care (MIMIC) and AmsterdamUMC databases, including basic demographics, vital signs, and laboratory results. We predicted the diagnosis of severe AKI in patients in the next 48 hours using machine-learning algorithms with the three databases. Then, we carried out real-time severe AKI prediction in the prospective validation study at our center for one year.

Participants: All patients included in three databases with uniform exclusion criteria.

Primary and secondary outcome measures: Effect evaluation index of prediction models.

Results: We included 58492 patients, and a total of 5257 (9.0%) patients met the definition of severe AKI. In the internal validation of the SHZJU-ICU and MIMIC databases, the best area under the receiver operating characteristic curve (AUROC) of the model was 0.86. The external validation results by AmsterdamUMC database were also satisfactory, with the best AUROC of 0.86. A total of 2532 patients were admitted to the center for prospective validation; 358 positive results were predicted, and 344 patients were diagnosed with severe AKI, with the best sensitivity of 0.72, the specificity of 0.80, and the AUROC of 0.84.

Conclusion: The prediction model of severe AKI exhibits promises as a clinical application based on dynamic vital signs and laboratory results of multi-center databases with prospective and external validation.

**Keywords**: Machine learning; Acute kidney injury; Real-time prospective validation; External validation.

**Strengths and limitations of this study**:

A prospective validation in machine learning of AKI research rather than other studies;

Three large database containing different national populations and regions;

Variable's sampling limited by the monitoring frequency of clinical data;

Differences in the samples proportion of three databases;

The dimensions of variables are not rich enough.

**Data sharing statement:** No data available.

## Introduction:

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalization and mortality[1]. AKI is divided into three types according to the various aetiologies: prerenal (renal hypoperfusion), intrarenal (vascular, glomerular, or tubulointerstitial lesions), and postrenal (urinary tract obstruction)[2]. Although nearly all diseases associated with ICU admission may cause AKI, acute tubular necrosis (ATN) and prerenal azotaemia are the most common causes[3].

All AKI diagnostic criteria including the latest Kidney Disease: Improving Global Outcomes (KDIGO) standard are currently based on the creatinine level and urine volume[4]. However, the increase in the creatinine level or decrease in the urine volume lags the onset of AKI[2]. Many studies have suggested that early diagnosis and treatment of reversible AKI can reduce mortality[5]. Therefore, the creatinine level and urine volume are not satisfactory to meet clinical diagnostic demands. Consequently, many

researchers have tried to develop an early warning model by analysing the risk factors for AKI[6].

Patient complications, such as diabetes, hypertension, cardiovascular disease, chronic liver disease, sepsis, and trauma, are identified as important risk factors for AKI[7]. The AKI prediction model and scoring system developed based on high-risk factors has gradually become the focus of research considering the lower clinical application threshold compared with that of new biomarkers[6]. Although most previous prediction models use the multiple logistic regression model, a variety of AKI prediction models based on machine learning have resulted in satisfactory outcomes[6]. Since the first AKI prediction model study based on artificial intelligence was published in 2016, researchers have built more than 20 published AKI prediction models successively by using local or multi-centre databases[6,8-14]. The results indicate that these models can predict the occurrence of AKI and the need for renal replacement therapy (RRT) within 24 or 48 hours, with accuracies ranging from 81% to 97%[6,15]. In addition, many studies have focused on subspecialized conditions, including cardiac surgery, trauma, and burns[14-16]. However, the common defect in these studies is the lack of external validation and prospective validation, which causes the prediction model to deviate from the clinical scenarios and limits extrapolation beyond the scope of the data.

In this study, we built models to predict AKI within 48 hours in critically ill patients by using transcontinental three databases. Then, we evaluated the clinical effect of the model through a one-year prospective validation at our centre.

## Methods:

### Study design and setting

We collected patients using three ICU databases and prospectively validated the models in our centre.

The first database was our centre general ICU database (SHZJU-ICU) of the Second Affiliated Hospital of Zhejiang University School of Medicine, an academic teaching hospital. Since its establishment in 2017, it has included 12000 ICU patients' data and is updated daily. The Medical Information Mart for Intensive Care (MIMIC) III database, the second one, is an open ICU database provided by the Massachusetts Institute of Technology and includes nearly 60000 ICU patients from North America[17]. Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23000 patients admitted to ICUs in parts of Europe[18]. The research flow chart is shown in Figure 1.

**Study definition**

In this study, the diagnosis of AKI was confirmed based on three stages according to the KDIGO criteria[4]. We defined the patients who met the KDIGO AKI II and III criteria as severe AKI groups and the others as negative groups. We excluded patients with lack of creatinine measurements during admission, patients with creatinine baseline more than 3.0 mg/dL at admission, patients who met severe AKI diagnosis within 24 hours, and patients who used RRT within 48 hours after admission[19]. In addition, we excluded pregnant women, patients younger than 14 years old, and patients hospitalized in the ICU for fewer than 48 hours. After the patient was admitted to ICU, we performed a prediction every 24 hours and recorded a prediction time. If the patient was diagnosed with severe AKI within 48 hours, the predictive time was defined as a positive predictive point, and the others were defined as a negative point. The study was evaluated and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine as study number 2019-078.

**Data Collection**

The variables included demographic data, vital signs, basic and primary diseases, laboratory results,

important operation records, and drug records. Comorbidity included hypertension, diabetes, cardiopathy, liver disease, and malignant tumors. The primary disease was the main cause of admission to the ICU following the ICD-10 codes. The vital signs and clinical laboratory results were transformed into different variables according to the average, variance, maximum, minimum, and final value before diagnosis. We use a method similar to the forward incremental method in the multivariate logic regression model, that is, the combination of embedded feature selection and forward addition for feature selection. First of all, all variables are trained in the model, then list by variables importance. variables are added to the model one by one according to the variable importance. a variable is retained if it causes the AUC growth to be greater than 0.01, otherwise delete it. We transformed the MIMIC and AmsterdamUMC databases according to our centre database structure, unifying the unit and diagnostic codes. We deleted variables missing more than 50%. Variables missing more than 30% but less than 50% are listed to clinicians who determine the potential correlation between these variables and AKI. We carry out multiple interpolation for these variables which clinicians require to be retained, and the others deleted. Variables missing less than 30% are fill in multiple interpolation. All Missing data between three databases and values included in the model shown in supplementation file Table S1.

**Model construction and external validation**

The ratio of the training and internal validation sets was 4:1. The SHZJU and MIMIC databases training sets were mixed into a new training set. There were more negative data than positive data, so we randomly sampled the negative datasets and constructed a new data subset with a sampling ratio of positive and negative data of 1:5 in model building in order to extract the importance variables. In the subsequent model validation, we adopted the original data set. We used multiple logistic regression, random forest, XGBoost, AdaBoost, LightGBoost, gradient boosting decision tree (GBDT), and debug

to assess the variables and model-related parameters by the fivefold cross-validation method. After the models were built, we used the SHZJU-ICU and MIMIC test sets for internal validation of the model and the AmsterdamUMC database for external validation. The most appropriate cut-off value was determined according to the K-S curve. The prediction model represents the results of each prediction with a probability between 0 to 1.0. We define results more than 0.4 as high-risk, that is, positive results, and the rest as negative results. Through internal validation and external validation, we calibrated the model by adjusting the super-parameters and using the Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. All model building and validation processes were performed in Python 3.6.

**Prospective validation**

The prospective research period was 2020.01.01 to 2020.12.31. We collected real-time data when patients were admitted to the ICU, transformed the data according to the requirements, and formed a complete sample for the prediction model after passing the integrity test. We had established a visualization scheme and allowed researchers to review the predictions daily. The daily prediction results were not publicly accessible during the study to avoid affecting clinicians' decisions, but the diagnosis results were available to the researchers as visual graphics. We sampled the 20% predicted data every month and deleted samples with more than 50% missing values to ensure data correctness. When a patient has the following conditions, AKI prediction system will end the patient's prospective prediction: A. a positive diagnosis; B. Transfer out of ICU or death with negative diagnosis. All diagnosis of severe AKI needs to be reviewed by two ICU attending physicians independently, and if the they have different opinion, the third one will be appealed.

**Statistical analysis:**

The population characteristics were reported as the medians and inter-quartile ranges (IQRs) for skewed data and the means and standard deviations for normally distributed data. The independent sample T-test was used for normally distributed data, and the rank-sum test was used for the rest. Dichotomous variables were assessed by the chi-square test, and a P value less than 0.05 was considered statistically significant. The non-normally distributed data were analyzed by exponential transformation and logarithmic transformation. The effect of the model was evaluated by parameters such as the AUROC, accuracy, specificity, and F1-score.

**Patient and Public Involvement:**

The information of cases in three databases was in a state of complete desensitization in the process of building the model. During the prospective study, all the patients signed an informed consent form at the beginning of admission to ICU. The real-time data discussed and used by only the study members, and were not made public during the study period. All data were anonymized before the authors accessed them for the purpose of this study. Therefore, patients' priorities, experience, and preferences will not affect the development of the research question and outcome measures. If necessary, we will inform patients of relevant research results by telephone.

**Results:**

According to the inclusion criteria and exclusion criteria, we selected 58492 patients from three databases who met the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36690 patients from the MIMIC database, and 15341 patients from the AmsterdamUMC database. A total of 5257 (9.0%) patients met the definition of severe AKI (11.8% in SHZJU-ICU, 7.6% in MIMIC, and 10.9% in AmsterdamUMC). The distributions of age and sex in the three centers were similar, but

the differences in race were large. Asian patients accounted for more than 99% in the SHZJU-ICU database, and only approximately 2.5% in the MIMIC database. White people accounted for more than 70% of the MIMIC database. In addition, patients from the MIMIC database had a higher incidence of the tumor, liver cirrhosis, diabetes, and hypertension. Patients in the AmsterdamUMC and SHZJU-ICU databases had a higher proportion of mechanical ventilation and overall survival rate. Severe AKI patients had longer ICU hospital stays and higher mortality. More details are presented in Table 1.

Table 1: Clinical Demographics and Outcomes in patients with or without severe acute kidney injury.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | MIMIC Negative (N=33879) | MIMIC Severe AKI (N=2811) | P-value | AmsterdamUMC Negative (N=13661) | AmsterdamUMC severe AKI (N=1680) | P-value |
|---|---|---|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 62.0 [45.0-76.0] | 64.0 [52.0-76.0] | <0.001 | 64.5 [54.5,74.5] | 64.3 [52.5,78.5] | 0.04 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 18976 (56.0) | 1524 (54.2) | 0.06 | 8920 (65.3) | 1054 (63.9) | 0.03 |
| **Race** | | | | | | | / | / | / |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 23796 (70.2) | 2005 (71.3) | 0.23 | / | / | / |
| **Black** | 1 (0.02) | 0 | 1 | 2455 (7.2) | 199 (7.1) | 0.45 | / | / | / |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 918 (2.7) | 70 (2.5) | 0.54 | / | / | / |
| **Other** | 0 | 1 (0.1) | 1 | 6710 (19.8) | 537 (19.1) | 0.34 | / | / | / |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 1.1 (0.41) | 1.1 (0.50) | 0.04 | 1.2 (0.35) | 1.1 (0.38) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (12) | 23 (12) | <0.001 | 18 (13) | 23 (10) | <0.001 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 6652 (19.6) | 752 (26.8) | <0.001 | / | / | / |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 1180 (3.5) | 426 (15.2) | <0.001 | / | / | / |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 4840 (14.3) | 336 (12.0) | 0.001 | / | / | / |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 7250 (21.4) | 789 (28.1) | <0.001 | / | / | / |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 16328 (48.2) | 1513 (53.8) | <0.001 | / | / | / |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 13009 (38.4) | 1503(53.5) | <0.001 | 7718(56.5) | 1248(74.3) | <0.001 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 12455 (36.7) | 1430 (50.8) | <0.001 | / | / | / |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 51.0 [28.5-110.3] | 103.4 [50.1-261.3] | <0.001 | 24.0 [19.8,63.5] | 142.0 [45.4,394.2] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 25648 (75.7) | 1356 (48.2) | <0.001 | 10942(80.1) | 1033 (61.5) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

There were significant differences in the important parameters of the variables among the different models (see Figure S1). However, the trend of the creatinine level in the past week was still an important variable, followed by urine volume, blood urea nitrogen level, temperature, and length of ICU stay. The cut-off value used to distinguish between a negative and positive prediction was determined by the K-S curve, with value of 0.423 (see Figure S2). The GBDT model had the best prediction effect in the test set, followed by XGBoost and LightGBoost. In the two central internal validation sets, the two best-performing machine learning algorithms with great AUROC are LightGBoost (SHZJU-ICU of 83.2%, MIMIC of 86.0%) and XGBoost (SHZJU-ICU 85.9%, MIMIC 85.6%), as detailed in Figure 2. Overall, the sensitivity (SHZJU-ICU 0.84, MIMIC 0.83) and the negative predictive value (SHZJU-ICU 0.90, MIMIC 0.90) of the predictive model were high, but the specificity was general (SHZJU-ICU 0.79, MIMIC 0.75), as shown in Table 2. In the external validation based on AmsterdamUMC database, the overall model validation effect results were satisfactory, and XGBoost had the best performance, with an AUROC of 0.84, as shown in Figure 2 and Table 2.

Table 2: Model validation results by three databases with machine learning algorithm

| *Model* | *AUROC* | *Accuracy* | *Sensitivity* | *Specificity* | *PPV* | *NPV* | *F1* |
|---|---|---|---|---|---|---|---|
| ***Internal validation with SHZJU-ICU database*** | | | | | | | |
| ***Logistic regression*** | *0.748* | *0.662* | *0.834* | *0.576* | *0.496* | *0.874* | *0.622* |
| ***LightGBoost*** | *0.832* | *0.741* | *0.839* | *0.692* | *0.576* | *0.896* | *0.683* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ***GBDT*** | *0.845* | *0.765* | *0.843* | *0.725* | *0.606* | *0.902* | *0.705* |
| ***AdaBoost*** | *0.806* | *0.721* | *0.824* | *0.67* | *0.555* | *0.884* | *0.663* |
| ***Random Forest*** | *0.821* | *0.763* | *0.71* | *0.789* | *0.627* | *0.845* | *0.666* |
| ***XGBoost*** | ***0.859*** | *0.779* | *0.81* | *0.763* | *0.631* | *0.889* | *0.709* |
| ***Internal validation with MIMIC database*** | | | | | | | |
| ***Logistic regression*** | *0.733* | *0.695* | *0.643* | *0.72* | *0.535* | *0.801* | *0.584* |
| ***LightGBoost*** | ***0.86*** | *0.768* | *0.822* | *0.741* | *0.613* | *0.893* | *0.702* |
| ***GBDT*** | *0.846* | *0.765* | *0.786* | *0.755* | *0.616* | *0.876* | *0.691* |
| ***AdaBoost*** | *0.837* | *0.732* | *0.831* | *0.683* | *0.567* | *0.89* | *0.674* |
| ***Random Forest*** | *0.832* | *0.738* | *0.791* | *0.712* | *0.578* | *0.872* | *0.668* |
| ***XGBoost*** | *0.856* | *0.758* | *0.833* | *0.721* | *0.598* | *0.895* | *0.695* |
| ***External validation with AmsterdamUMC database*** | | | | | | | |
| ***Logistic regression*** | *0.704* | *0.767* | *0.516* | *0.893* | *0.706* | *0.787* | *0.596* |
| ***LightGBoost*** | *0.859* | *0.763* | *0.827* | *0.731* | *0.606* | *0.894* | *0.7* |
| ***GBDT*** | *0.861* | *0.764* | *0.84* | *0.727* | *0.606* | *0.901* | *0.704* |
| ***AdaBoost*** | *0.85* | *0.755* | *0.813* | *0.726* | *0.597* | *0.886* | *0.689* |
| ***Random Forest*** | *0.82* | *0.743* | *0.77* | *0.729* | *0.587* | *0.864* | *0.666* |
| ***XGBoost*** | ***0.865*** | *0.75* | *0.873* | *0.688* | *0.584* | *0.916* | *0.7* |
| ***Prospective validation with SHZJU-ICU*** | | | | | | | |
| ***Logistic regression*** | *0.758* | *0.772* | *0.648* | *0.834* | *0.662* | *0.826* | *0.655* |
| ***LightGBoost*** | *0.819* | *0.796* | *0.596* | *0.895* | *0.74* | *0.816* | *0.66* |
| ***GBDT*** | *0.827* | *0.781* | *0.706* | *0.818* | *0.66* | *0.848* | *0.683* |
| ***AdaBoost*** | *0.808* | *0.766* | *0.686* | *0.805* | *0.638* | *0.837* | *0.661* |
| ***Random Forest*** | *0.804* | *0.755* | *0.715* | *0.775* | *0.613* | *0.845* | *0.66* |
| ***XGBoost*** | ***0.841*** | *0.779* | *0.724* | *0.807* | *0.652* | *0.854* | *0.686* |

*AUROC*: Area Under Receiver Operating Characteristic; *PPV*: Positive Predictive Value; *NPV*: Negative Predictive Value; *SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *GBDT*: Gradient Boosted Decision Tree.

According to the inclusion and exclusion criteria, we delete 267 patients among 94 patients with creatinine baseline more than 3.0 mg/dL at admission, 39 patients met severe AKI diagnosis within 24 hours, and 26 patients who used RRT within 48 hours after admission, 108 patients hospitalized in the ICU for fewer than 48 hours. A total of 2532 patients were admitted to our centre for prospective validation, and the prediction model made 16858 times predictions. In the prospective cohort, there was no significant difference in age, gender, baseline creatinine and urea nitrogen, and complications. The

proportion of mechanical ventilation and the ICU stay time in AKI patients were longer with higher mortality. Above all, there was no significant difference between the prospective and the retrospective cohort. More detail sees in supplementation file Table S2. In the end, 358 positive results were predicted, and the rest were negative results. There are 344 patients with severe AKI were diagnosed and the prediction accuracy was 83.5%. The model with the highest area under the curve was XGBoost, 0.84 with the best sensitivity of 0.72, the specificity of 0.80. The results of the prospective study are similar to those of the external validation of the model, and are relatively stable. More detail is presented in Figure 3 and Table 2.

**Discussion:**

In this study, we built predictive models by machine learning to predict the incidence of severe AKI with three databases in different regions and in the next 48 hours. After internal and external validation, prospective validation over one year was carried out to verify the model effects. The three databases come from three countries that are in Asia, Europe, and North America, which proves that the model is universal to some extent.

Despite the huge amount of data, many databases are still not suitable for prospective research because they are not updated promptly. Tomasev, N and colleagues have provided research on AKI prediction models with a large amount of data[10]. The study covered 703782 adult patients with 6 billion individual items, including 620000 elements. In this study, a depth neural network model was used for real-time prediction. A total of 55.8% of severe AKI patients were predicted within the first 48 hours, although each accurate prediction was accompanied by two mispredictions[10]. This study provided a new scheme for real-time prediction and indicated that we should prospectively evaluate and

independently validate models to explore their effectiveness. In a prospective study, Flechet, M et al compared an AKI prediction model with clinicians in 252 patients and found that the clinical effect of the random forest model for predicting AKI-II/III was equivalent to that of clinicians. Our prediction model graphical visualization of the model was installed in the centre's database for better usage. In addition, our database is updated daily to achieve daily predictions and present the results to researchers. In the prospective validation of our study, the stability of the prediction model confirmed its promise, which provides a basis for future research.

There are many studies of artificial intelligence for predicting the occurrence of AKI, but most of them are single-centre studies, and the extrapolation effect has been controversial. In 2016, Koyner, J, and his colleague published the first study of an AKI prediction model based on multi-centre data. In all 202961 patients, 17541 (8.6%) had AKI, 4251 (3.5%) had AKI-II, and 1242 (0.6%) had AKI-III. A multivariate logistic regression model was used to predict AKI in this study with an AUROC of 0.74 (95% CI=0.74). With the classification of AKI, the AUROC of the prediction model gradually increased to 0.84[8]. Subsequently, the study team used a new machine learning algorithm, to build a more accurate model to predict the occurrence of AKI-II, with an AUROC of 0.9 within 24 hours and 0.87 within 48 hours[19]. Recently, the research team included data from two other centres, namely, LUMC (N=200613) and NUS (N=246895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85 to 0.86, suggesting that the artificial intelligence model has stable predictive ability[12]. This series of studies included many data points, suggesting the feasibility of artificial intelligence in the diagnosis of AKI, but the proportion of positive patients (3.5%) and ICU patients (30%) was too low to properly predict AKI. Our research is similar to the above. The SAHZJU-ICU database is a single-centre database representing south-eastern China, and the MIMIC database is a well-known open ICU database in the

United States. The AmsterdamUMC database is a public database located in Europe. The population structure and diseases in the three databases are complete but different in the distribution of complications and race. Therefore, it provides a prediction model with unparalleled stability compared with other studies.

Limitations:

This retrospective multi-centre study was unable to carry out more clinical feature mining and comparison because of different data structures. The differences between the three databases partly reflect some demographic differences between Europe, the United States and China, resulting in a decline in the accuracy of the prediction model. There are some differences in the number of patients included in the three databases, which may affect the choice of variables. As a result of the study design, we deleted patients with ICU hospitalization of less than 48 hours, which may result in the exclusion of most relatively mild patients and may reduce false positives. Second, in the prospective data study in 2020, there may be deviations in the inclusion of patients in the centre, thus affecting the interpretation of the follow-up prospective results. Finally, given the low incidence of severe AKI and the great difference in the proportion of positive and negative samples, the data may be accidental. Our model seems to be superior to diagnostic non-AKI patients rather than AKI because of the proportion of positive data that we include. In the retrospective study, we reduced the proportion of negative data by randomization but retain all date in prospective phase with the sensitivity decreases.

**Conclusion:**

Based on databases of patients of different races from different countries, we constructed stable

machine learning models to predict the occurrence of AKI in the next 48 hours. Prospective validation through the implementation of an updated local database is an effective exploration of further research.

Data sharing statement: No data available.

Ethics approval: Our research was approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

Competing interests for authors: There may be some potential competing interests between our research and "HealSci Technology Co. Beijing.". In this study, we jointly complete the work including data integration and transcoding in model construction and model visualization projects based on contract.

**Reference:**

1. Ronco, C.; Bellomo, R.; Kellum, J.A. Acute kidney injury. *The Lancet* **2019**, *394*, 1949-1964, doi:10.1016/s0140-6736(19)32563-2.
2. Pakula, A.M.; Skinner, R.A. Acute Kidney Injury in the Critically Ill Patient: A Current Review of the Literature. *J Intensive Care Med* **2016**, *31*, 319-324, doi:10.1177/0885066615575699.
3. Bellomo, R.; Kellum, J.A.; Ronco, C. Acute kidney injury. *The Lancet* **2012**, *380*, 756-766, doi:10.1016/s0140-6736(11)61454-2.
4. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* **2012**, *120*, c179-184, doi:10.1159/000339789.
5. Coelho, S.; Cabral, G.; Lopes, J.A.; Jacinto, A. Renal regeneration after acute kidney injury. *Nephrology (Carlton)* **2018**, *23*, 805-814, doi:10.1111/nep.13256.

6. Wilson, F.P.; Greenberg, J.H. Acute Kidney Injury in Real Time: Prediction, Alerts, and Clinical Decision Support. *Nephron* **2018**, *140*, 116-119, doi:10.1159/000492064.

7. Wu, Y.; Hao, W.; Chen, Y.; Chen, S.; Liu, W.; Yu, F.; Hu, W.; Liang, X. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* **2020**, *42*, 1127-1134, doi:10.1080/0886022X.2020.1843491.

8. Koyner, J.L.; Adhikari, R.; Edelson, D.P.; Churpek, M.M. Development of a Multicenter Ward-Based AKI Prediction Model. *Clin J Am Soc Nephrol* **2016**, *11*, 1935-1943, doi:10.2215/CJN.00280116.

9. Low, S.; Vathsala, A.; Murali, T.M.; Pang, L.; MacLaren, G.; Ng, W.Y.; Haroon, S.; Mukhopadhyay, A.; Lim, S.L.; Tan, B.H., et al. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* **2019**, *20*, 32, doi:10.1186/s12882-019-1206-4.

10. Tomasev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I., et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116-119, doi:10.1038/s41586-019-1390-1.

11. Flechet, M.; Falini, S.; Bonetti, C.; Güiza, F.; Schetz, M.; Van den Berghe, G.; Meyfroidt, G. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Critical Care* **2019**, *23*, 282, doi:10.1186/s13054-019-2563-x.

12. Churpek, M.M.; Carey, K.A.; Edelson, D.P.; Singh, T.; Astor, B.C.; Gilbert, E.R.; Winslow, C.; Shah, N.; Afshar, M.; Koyner, J.L. Internal and External Validation of a

Machine Learning Risk Score for Acute Kidney Injury. *JAMA Netw Open* **2020**, *3*, e2012892, doi:10.1001/jamanetworkopen.2020.12892.

13. Yuan, Q.; Zhang, H.; Deng, T.; Tang, S.; Yuan, X.; Tang, W.; Xie, Y.; Ge, H.; Wang, X.; Zhou, Q., et al. Role of Artificial Intelligence in Kidney Disease. *Int J Med Sci* **2020**, *17*, 970-984, doi:10.7150/ijms.42078.

14. Li, Y.; Xu, J.; Wang, Y.; Zhang, Y.; Jiang, W.; Shen, B.; Ding, X. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* **2020**, *43*, 752-761, doi:10.1002/clc.23377.

15. De Vlieger, G.; Kashani, K.; Meyfroidt, G. Artificial intelligence to guide management of acute kidney injury in the ICU: a narrative review. *Curr Opin Crit Care* **2020**, *26*, 563-573, doi:10.1097/MCC.0000000000000775.

16. James, M.T.; Pannu, N.; Hemmelgarn, B.R.; Austin, P.C.; Tan, Z.; McArthur, E.; Manns, B.J.; Tonelli, M.; Wald, R.; Quinn, R.R., et al. Derivation and External Validation of Prediction Models for Advanced Chronic Kidney Disease Following Acute Kidney Injury. *JAMA* **2017**, *318*, 1787-1797, doi:10.1001/jama.2017.16326.

17. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.W.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.A.-O.; Celi, L.A.; Mark, R.A.-O. MIMIC-III, a freely accessible critical care database.

18. Thoral, P.J.; Peppink, J.M.; Driessen, R.H.; Sijbrands, E.J.G.; Kompanje, E.J.O.; Kaplan, L.; Bailey, H.; Kesecioglu, J.; Cecconi, M.; Churpek, M., et al. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European

Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example. *Crit Care Med* **2021**, 10.1097/CCM.0000000000004916, doi:10.1097/CCM.0000000000004916.

19. Koyner, J.L.; Carey, K.A.; Edelson, D.P.; Churpek, M.M. The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit Care Med* **2018**, *46*, 1070-1077, doi:10.1097/CCM.0000000000003123.

**Footnotes**

Contributions: Conceptualization, QL and MH; Methodology, YX; Software, YX; Validation, YZ, and YX.; Formal Analysis, JC and XC; Resources, MH; Data Curation, QL; Writing – Original Draft Preparation, QL; Writing – Review & Editing, MH; Visualization, YX; Supervision, YZ and QL; Project Administration, QL;

**Funding**: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest: There may be some potential competing interests between our research and "HealSci Technology Co. Beijing.". In this study, we jointly complete the work including data integration and transcoding in model construction and model visualization projects based on contract.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Acknowledgements: Thank my colleagues in general ICU for cooperating with us in our prospective study. Thanks to two public databases namely MIMIC and AmsterdamUMC for providing important

data. Thanks for all patient advisers.

**Supplemental Figure:**

File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.

File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.

File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

**Figure Legend:**

Figure 1: The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.

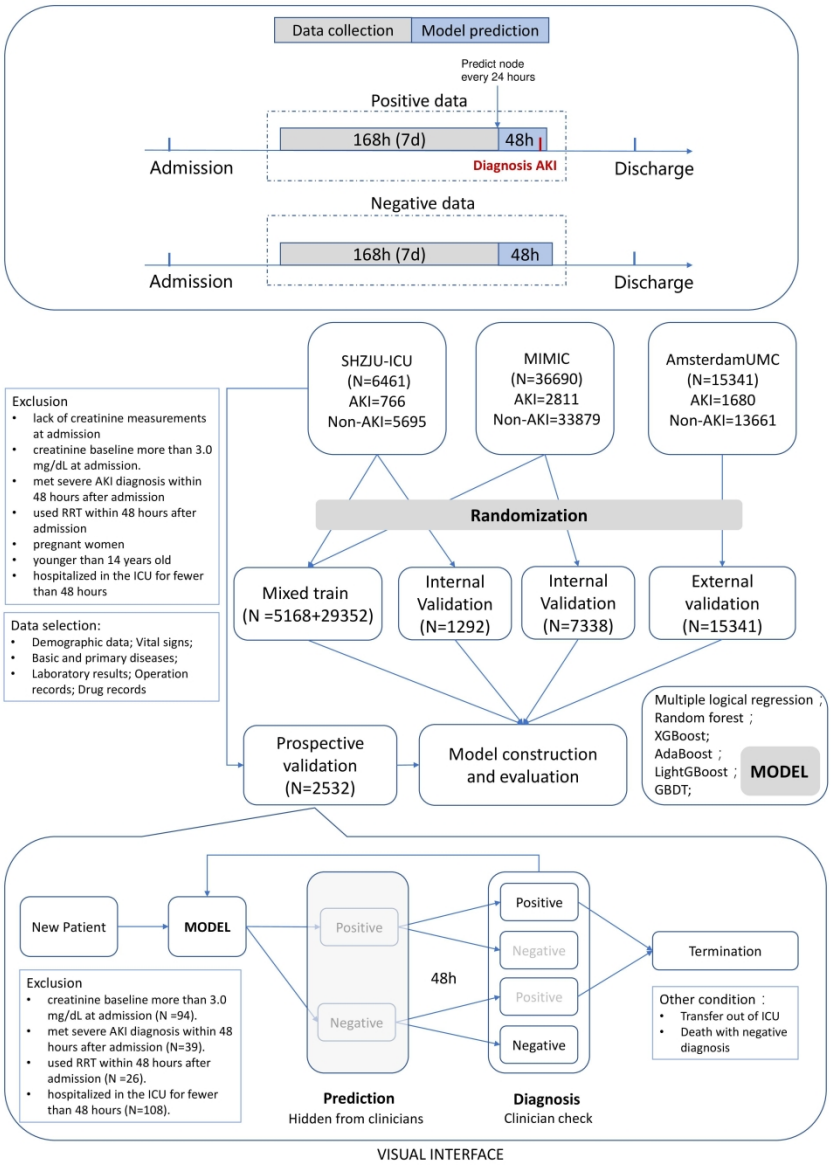



Figure 1: The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the MIMIC database and incorporated into the AmsterdamUMC database for external validation. We carried out a one-year prospective validation through the database of the centre.

338x451mm (300 x 300 DPI)

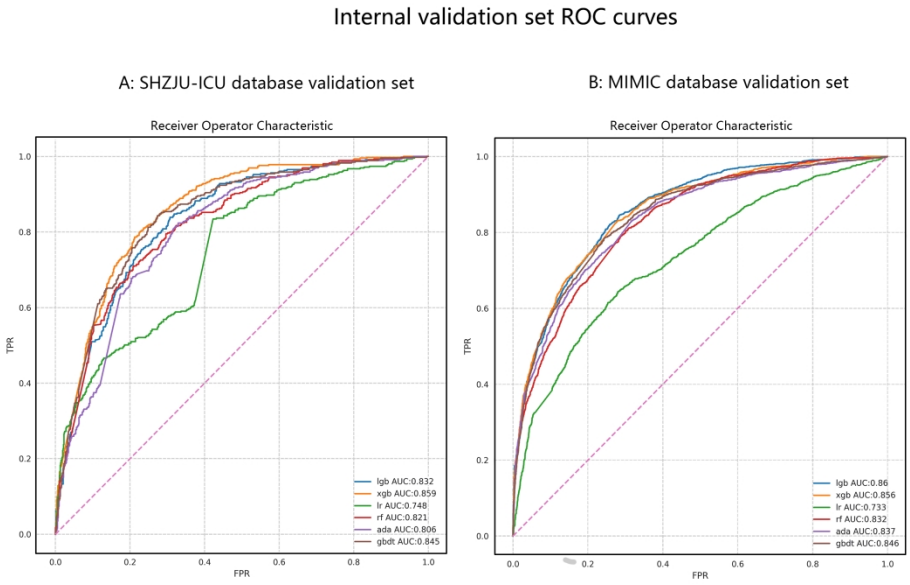



Figure 2: The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database.

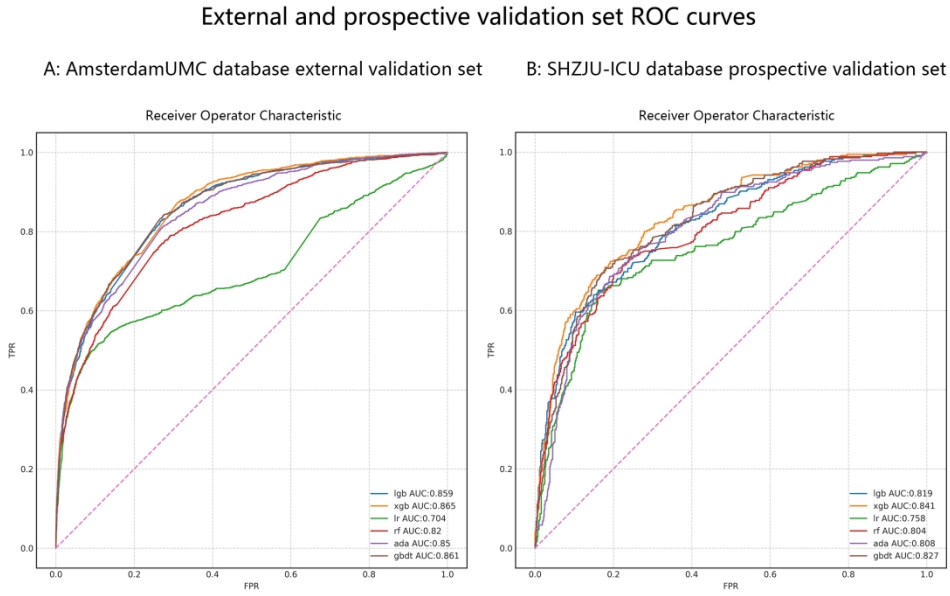
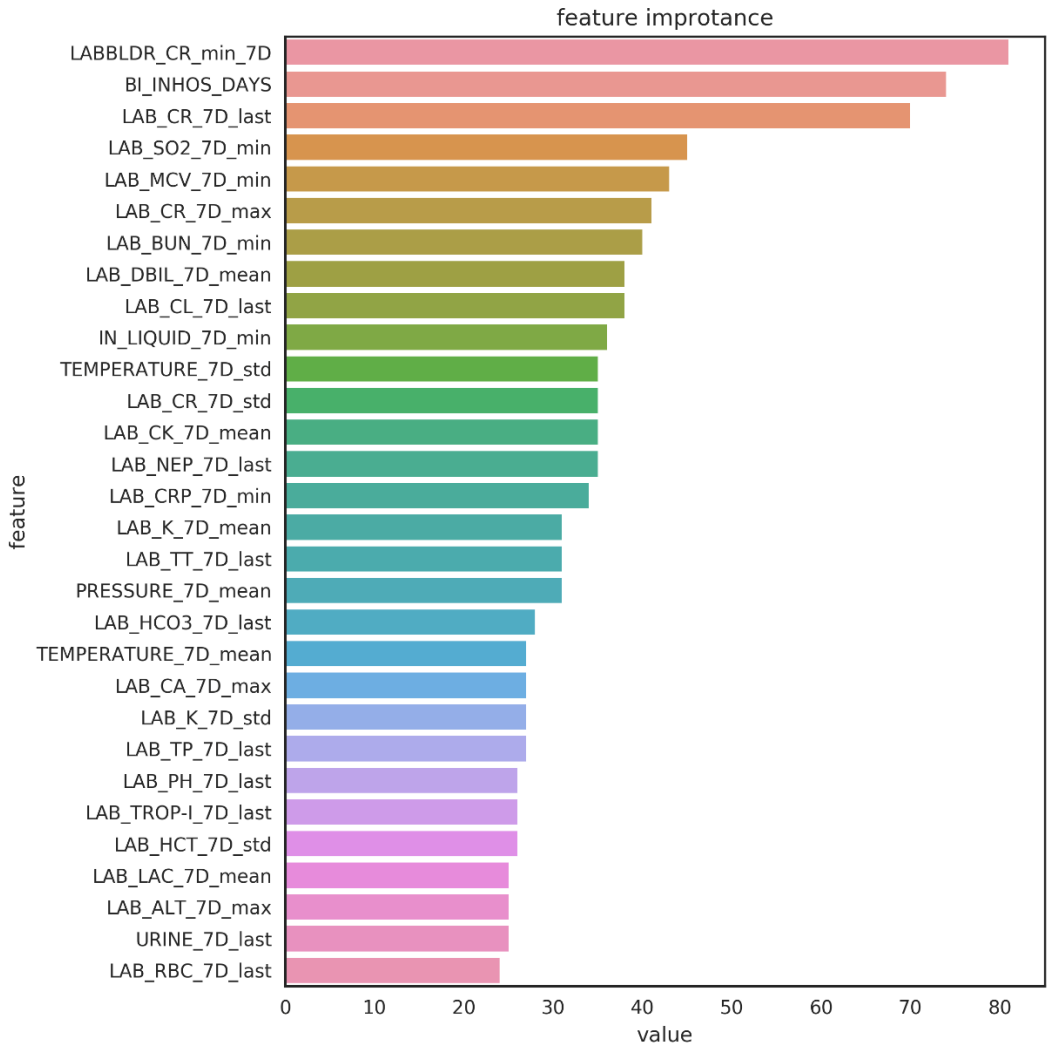



Figure 3: The AUROC curve of the external validation set of the AmsterdamUMC database for 3A and the prospective validation in our center for 3B.

**Supplemental Figure:**

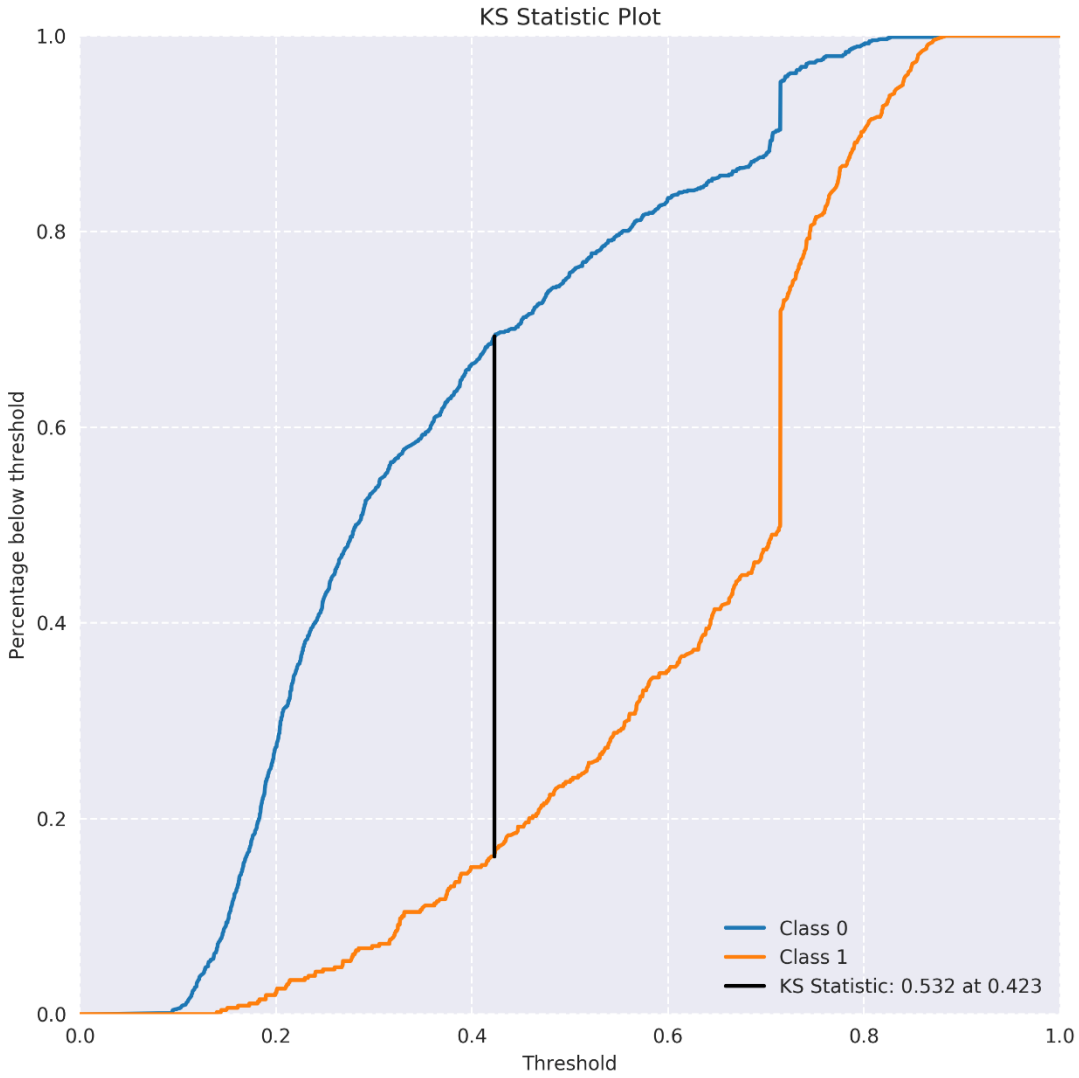
File name: **Figure S1.tif**

Title and Description of data: The variable importance parameter histogram of early diagnosis prediction models of severe AKI uses a multiple logistic regression algorithm and five machine learning algorithms.




File name: **Figure S2.tif**

Title and Description of data: The K-S curve used to determine the cut-off value of positive and negative diagnosis of early diagnosis prediction models of severe AKI.



File name: **Table S1.docx**

Title and of Description data: Missing data between three databases in the study and values included in the model.

| | SHZJU-ICU | MIMIC | AmsterdamUMC | Including in Model |
|---|---|---|---|---|
| **Age** | √ | √ | √ | √ |
| **Gender** | √ | √ | √ | √ |
| **Race** | √ | √ | × | × |
| **Primary disease** | √ | √ | × | × |
| **Comorbidity** | √ | √ | × | × |
| **Ventilation** | √ | √ | √ | √ |
| **Operation** | √ | √ | × | × |
| **ICU hours** | √ | √ | √ | √ |
| **Survived** | √ | √ | √ | √ |
| **Vital sign** | | | | |
| **Temperature** | √ | √ | √ | √ |
| **Systolic pressure** | √ | √ | √ | √ |
| **Diastolic pressure** | √ | √ | √ | √ |
| **Respiratory rate** | √ | √ | √ | √ |
| **Heart rate** | √ | √ | √ | √ |
| **Oxygen saturation** | √ | √ | √ | √ |
| **Urine** | √ | √ | √ | √ |
| **GCS score** | √ | × | × | × |
| **Laboratory results** | | | | |
| **Scr** | √ | √ | √ | √ |
| **BUN** | √ | √ | √ | √ |
| **RBC** | √ | √ | √ | √ |
| **MCV** | √ | √ | √ | √ |
| **Hb** | √ | √ | √ | √ |
| **WBC** | √ | √ | √ | √ |
| **NEUT** | √ | √ | √ | √ |
| **PLT** | √ | √ | √ | √ |
| **HCV** | √ | √ | × | × |
| **TBLB** | √ | √ | √ | √ |
| **DBLB** | √ | √ | √ | √ |
| **IBLB** | √ | √ | √ | √ |
| **CRP** | √ | √ | √ | √ |
| **PCT** | √ | × | × | × |
| **Serum kalium** | √ | √ | √ | √ |
| **Serum natrium** | √ | √ | √ | √ |
| **Serum chlorine** | √ | √ | √ | √ |

| | | | | |
|---|---|---|---|---|
| **PT** | √ | √ | √ | √ |
| **APTT** | √ | √ | √ | √ |
| **INR** | √ | √ | √ | √ |
| **CK** | √ | √ | √ | √ |
| **LDH** | √ | √ | √ | √ |
| **Troponin** | √ | √ | √ | √ |
| **Blood glucose** | √ | √ | √ | √ |
| **PH** | √ | √ | √ | √ |
| **Lactic acid** | √ | √ | √ | √ |
| **Anion gap** | √ | √ | √ | √ |
| **ABC** | √ | √ | √ | √ |
| **SBC** | √ | √ | √ | √ |
| **$PaCO_2$** | √ | √ | √ | √ |
| **$PaO_2$** | √ | √ | √ | √ |

SHZJU-ICU: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; MIMIC: Medical Information Mart for Intensive Care; APTT: Activated partial thromboplastin time; INR: International normalized ratio; LDH: Lactate dehydrogenase; $PaCO_2$: Partial pressure of carbon dioxide; $PaO_2$ Partial pressure of oxygen; MCV: Mean corpuscular volume; PH: Potential of hydrogen; WBC: White blood cell count; NEUT: Proportion of neutrophils; CK: Creatine kinase; Scr: Serum Creatinine; BUN: Blood urea nitrogen; RBC: Red blood cell count; Hb: Hemoglobin; PT: Prothrombin time; TBLB: Total bilirubin; DBLB: Direct bilirubin; IBLB: Indirect bilirubin; CRP: C-reactive protein; HCV: Hematocrit value; PCT: Procalcitonin; ABC: Actual Bicarbonate Radical; SBC: Standard Bicarbonate Radical; PLT: Platelet count

File name: **Table S2.docx**

Title and of Description data: Clinical Demographics and Outcomes of SHZJU-ICU patients in model building set and prospective validation.

| | SHZJU-ICU Negative (N=5695) | SHZJU-ICU Severe AKI (N=766) | P-value | SHZJU-ICU Prospective Negative (N=2188) | SHZJU-ICU Prospective Positive (N=344) | P-value |
|---|---|---|---|---|---|---|
| **Age, median [IQR]** | 60.5 [48.8-70.1] | 62.1 [50.1-72.5] | 0.004 | 61.3 [47.5-73.4] | 62.1 [51.2-70.5] | 0.03 |
| **Gender, male (%)** | 3561 (64.9) | 438 (59.8) | 0.007 | 1365 (62.4) | 207 (60.2) | 0.438 |
| **Race** | | | | | | |
| **White** | 3 (0.05) | 1 (0.1) | 1 | 2 (0.001) | 0 | 1 |
| **Black** | 1 (0.02) | 0 | 1 | 0 | 0 | 1 |
| **Asian** | 5691(99.9) | 764 (99.7) | 1 | 2185 (99.9) | 344(100) | 1 |
| **Other** | 0 | 1 (0.1) | 1 | 1 (0.001) | 0 | 1 |
| **Admission Scr (mg/dL), mean (sd)** | 0.9 (0.34) | 1.0 (0.40) | <0.001 | 0.8 (0.32) | 0.9 (0.35) | <0.001 |
| **Admission BUN (mg/dL), mean (sd)** | 18 (11) | 21 (9) | <0.001 | 17 (11) | 20 (10) | 0.007 |
| **Cancer, n (%)** | 644 (11.3) | 48 (6.3) | <0.001 | 230 (10.5) | 26 (7.6) | 0.102 |
| **Cirrhosis, n (%)** | 43 (0.8) | 6 (0.8) | 0.89 | 15 (0.7) | 3 (0.8) | 0.726 |
| **Cardiopathy, n (%)** | 231 (4.1) | 42 (5.5) | 0.08 | 98 (4.5) | 19 (5.7) | 0.403 |
| **Diabetes, n (%)** | 305 (5.4) | 63 (8.2) | 0.002 | 142 (6.5) | 29 (8.6) | 0.203 |
| **Hypertension, n (%)** | 889 (15.6) | 115 (15.0) | 0.70 | 324(14.8) | 56 (16.4) | 0.466 |
| **Ventilation, n (%)** | 2569 (45.1) | 592 (77.3) | <0.001 | 1102 (50.4) | 239 (69.5) | <0.001 |
| **Operation, n (%)** | 3361 (59.1) | 373 (48.7) | <0.001 | 1234(56.4) | 160 (46.7) | 0.001 |
| **ICU Hours, median [IQR]** | 50.8 [23.0-139.5] | 164.4 [70.3-328.6] | <0.001 | 58.7 [34.2-160.4] | 145.6 [68.5-314.8] | <0.001 |
| **Survived, n (%)** | 4997 (87.7) | 380 (49.6) | <0.001 | 1868 (85.4) | 178 (51.9) | <0.001 |

*SHZJU-ICU*: The general ICU database of the second affiliated hospital of Zhejiang university school of medicine; *MIMIC*: Medical Information Mart for Intensive Care; *IQR*: Inter-Quartile Range; *ICU*: Intensive Care Unit; *AKI*: Acute Kidney Injury. *Scr*: serum creatinine; *BUN*: blood urea nitrogen.

# TRIPOD Checklist: Prediction Model Development and Validation

| 1Section/Topic | Item | | Checklist Item | Page |
|---|---|---|---|---|
| **Title and abstract** | | | | |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Yes; P1 |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Yes; P2 |
| **Introduction** | | | | |
| Background and objectives | 3a | D;V | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Yes; P3 |
| | 3b | D;V | Specify the objectives, including whether the study describes the development or validation of the model or both. | Yes; P4, L76-79 |
| **Methods** | | | | |
| Source of data | 4a | D;V | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Yes; P4, L83-92 |
| | 4b | D;V | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Yes; P4, L83-100 |
| Participants | 5a | D;V | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Yes |
| | 5b | D;V | Describe eligibility criteria for participants. | Yes; P4, L103 |
| | 5c | D;V | Give details of treatments received, if relevant. | NA |
| Outcome | 6a | D;V | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Yes; P6, L111 |
| | 6b | D;V | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | 7a | D;V | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Yes; P6, L116 |
| | 7b | D;V | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | 8 | D;V | Explain how the study size was arrived at. | NA |
| Missing data | 9 | D;V | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Yes; P6, L127 |
| Statistical analysis methods | 10a | D | Describe how predictors were handled in the analyses. | Yes; P6, L125 |
| | 10b | D | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | Yes; P6, L130-138 |
| | 10c | V | For validation, describe how the predictions were calculated. | Yes; P6, L139 |
| | 10d | D;V | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Yes; P7, L159 |
| | 10e | V | Describe any model updating (e.g., recalibration) arising from the validation, if done. | Yes; P7, L137 |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done. | Yes; P6, L111 |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Yes; P6, L132 |
| **Results** | | | | |
| Participants | 13a | D;V | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Yes; Figure 1 |
| | 13b | D;V | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Yes, P8, L168-180 and SupTable2 |
| | 13c | V | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Yes, SupTable3 |
| Model development | 14a | D | Specify the number of participants and outcome events in each analysis. | Yes, P8, L168 |
| | 14b | D | If done, report the unadjusted association between each candidate predictor and outcome. | Yes, SupTable1 |
| Model specification | 15a | D | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | Yes. |
| | 15b | D | Explain how to the use the prediction model. | Yes, P9, L196 |
| Model performance | 16 | D;V | Report performance measures (with CIs) for the prediction model. | Yes, P9, L190 |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). | NA |
| **Discussion** | | | | |
| Limitations | 18 | D;V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Yes, P12, L246 |
| Interpretation | 19a | V | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Yes, P11, L237 |
| | 19b | D;V | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Yes, P10 L210-245 |
| Implications | 20 | D;V | Discuss the potential clinical use of the model and implications for future research. | Yes, P12, L254 |
| **Other information** | | | | |
| Supplementary information | 21 | D;V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Yes |
| Funding | 22 | D;V | Give the source of funding and the role of the funders for the present study. | NA |

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.