BMJ Open

Severe acute kidney injury predicting model based on transcontinental databases: a single-centre prospective study

Qiqiang Liang, Yongfeng Xu, Yu Zhou, Xinyi Chen, Juan Chen, Man Huang

General intensive care unit, Zhejiang University School of Medicine Second Affiliated Hospital, Hangzhou, Zhejiang, China

Correspondence to
Dr Man Huang;
huangman@zju.edu.cn

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 1 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 58 492 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 centre 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 multicentre databases with prospective and external validation.

INTRODUCTION

Acute kidney injury (AKI), as a common clinical complication in the intensive care unit (ICU), significantly increases the duration of hospitalisation 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 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 multicentre 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 subspecialised 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 1-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 12 000 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 60 000 ICU patients from North America.[17] Lastly, the AmsterdamUMC database is an available European ICU database with health data related to 23 000 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 hospitalised 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 tumours. 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 online supplemental 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 and 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
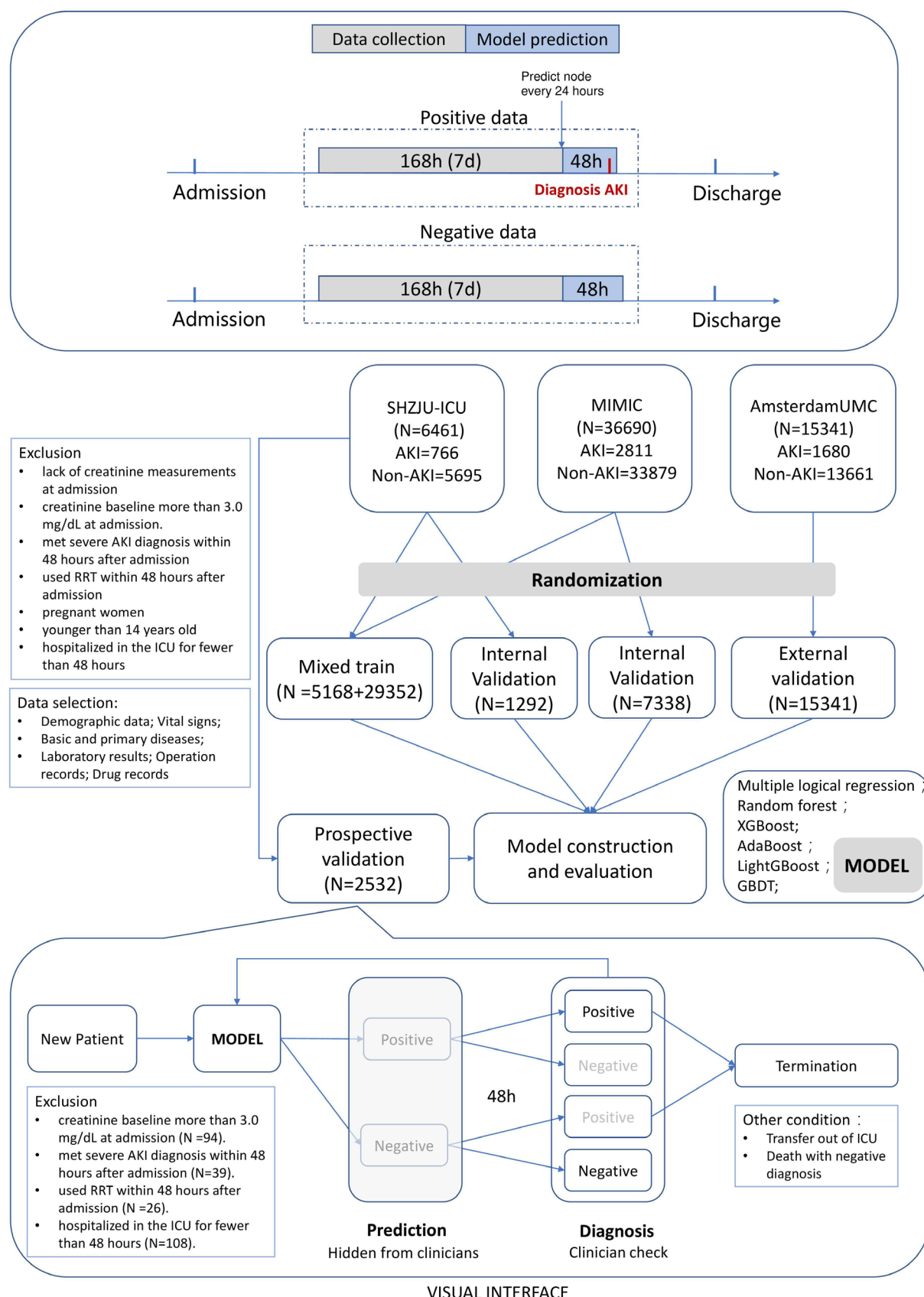
Figure 1 The research flow chart. The data collection time interval of the study was 7 days before diagnosis and the prediction interval was 48 hours. The early prediction model of AKI diagnosis was constructed and verified by our database and the mimic database and incorporated into the AmsterdamUMC database for external validation. We carried out a 1-year prospective validation through the database of the centre. AKI, acute kidney injury; GBDT, gradient boosting decision tree; ICU, intensive care unit; MIMIC, Medical Information Mart for Intensive Care; RRT, renal replacement therapy.

Platt calibration algorithm and compared the calibration effect by drawing a reliability diagram. All model building and validation processes were performed in Python V.3.6.

Prospective validation

The prospective research period was 1 January 2020–31 December 2020. 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 visualisation 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 IQRs for skewed data and the means and SD 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 χ^2 test, and a p value less than 0.05 was considered statistically significant. The non-normally distributed data were analysed by exponential transformation and logarithmic transformation. The effect of the model was evaluated by parameters such as the area under the receiver operating characteristic (AUROC), accuracy, specificity and F1-score.

Patient and public involvement

The information of cases in three databases was in a state of complete desensitisation 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 anonymised 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 58 492 patients from three databases who met

the requirements of the study, including 6461 patients from the SHZJU-ICU database, 36 690 patients from the MIMIC database and 15 341 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 centres 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 tumour, 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.

There were significant differences in the important parameters of the variables among the different models (see online supplemental 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 online supplemental 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.

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 hospitalised 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 16 858 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

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=33 879)	MIMIC Severe AKI (N=2811)	P value	AmsterdamUMC negative (N=13 661)	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	18 976 (56.0)	1524 (54.2)	0.06	8920 (65.3)	1054 (63.9)	0.03
Race			1						/
White	3 (0.05)	1 (0.1)	1	23 796 (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	16 328 (48.2)	1513 (53.8)	<0.001	/	/	/
Ventilation, n (%)	2569 (45.1)	592 (77.3)	<0.001	13 009 (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	12 455 (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	25 648 (75.7)	1356 (48.2)	<0.001	10 942 (80.1)	1033 (61.5)	<0.001

AKI, acute kidney injury; BUN, blood urea nitrogen; ICU, intensive care unit; MIMIC, Medical Information Mart for Intensive Care; Scr, serum creatinine.
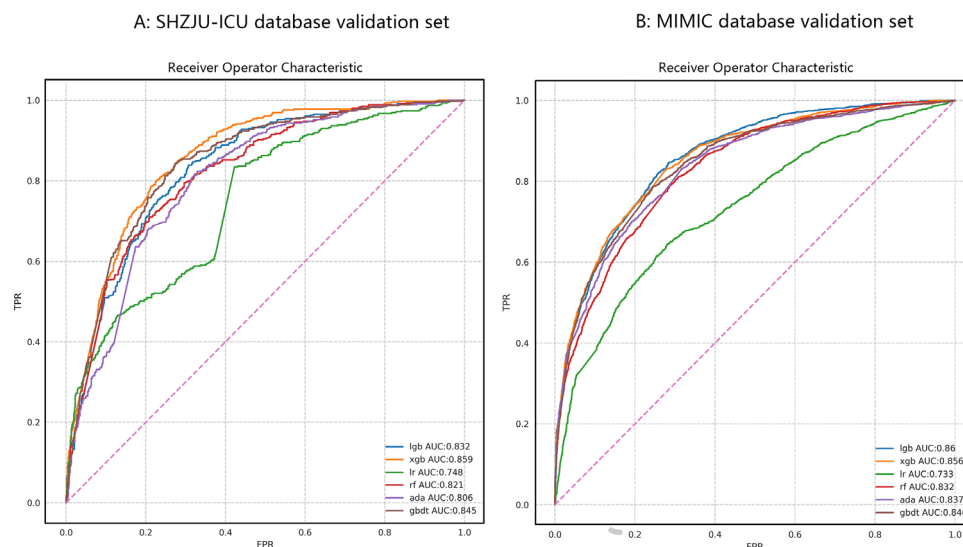
Internal validation set ROC curves

A: SHZJU-ICU database validation set

B: MIMIC database validation set

Figure 2 The AUROC curve of the internal validation set of the SHZJU-ICU database and the MIMIC database. AUROC, area under the receiver operating characteristic curve; ICU, intensive care unit; FPR, False Positive Rate; MIMIC, Medical Information Mart for Intensive Care; ROC, receiver operating characteristic; TPR, True Positive Rate.

in sonline supplemental 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 1 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. Tomašev *et al* have provided research on AKI prediction models with a large amount of data.[10] The study covered 703 782 adult patients with 6 billion individual items, including 620 000 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 *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 visualisation 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. Koyner *et al* published the first study of an AKI prediction model based on multicentre data. In all 2 02 961 patients, 17 541 (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. 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=2 00 613) and NUS (N=2 46 895), to externally validate the AKI-II prediction model with AUROCs reaching 0.85–0.86, 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

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; GBDT, Gradient Boosted Decision Tree; MIMIC, Medical Information Mart for Intensive Care; NPV, negative predictive value; PPV, positive predictive value.

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 USA. 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 multicentre 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 USA 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 hospitalisation 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 randomisation but retain all date in prospective phase with the sensitivity decreases.

External and prospective validation set ROC curves

A: AmsterdamUMC database external validation set   B: SHZJU-ICU database prospective validation set
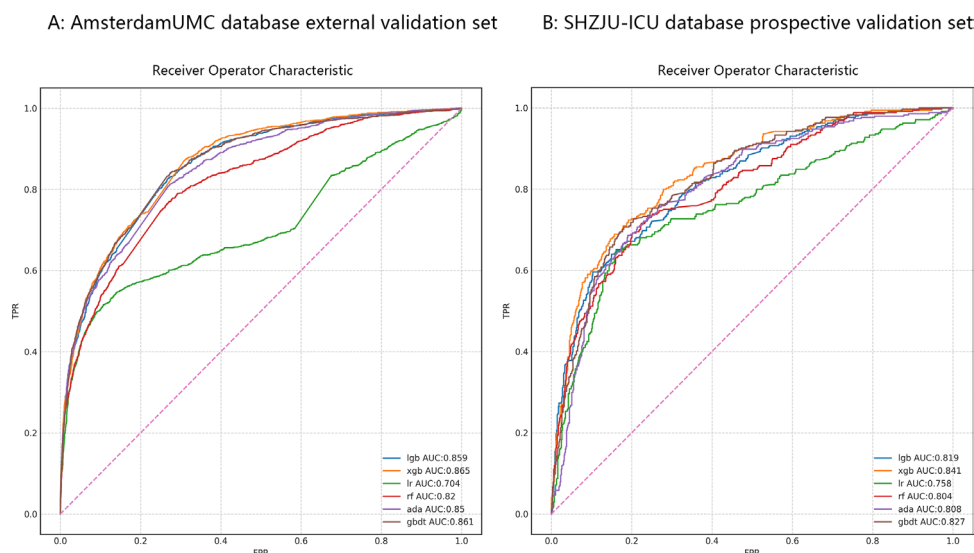

Figure 3 The AUROC curve of the external validation set of the AmsterdamUMC database for (A) and the prospective validation in our centre for (B). AUROC, area under the receiver operating characteristic curve; ICU, intensive care unit; ROC, receiver operating characteristic.

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.

Competing interests 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 visualisation projects based on contract.

Patient consent for publication Not applicable.

Ethics approval 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.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iD
Man Huang http://orcid.org/0000-0002-0842-2254

REFERENCES

1 Ronco C, Bellomo R, Kellum JA. Acute kidney injury. *Lancet* 2019;394:1949–64.
2 Pakula AM, Skinner RA. Acute kidney injury in the critically ill patient: a current review of the literature. *J Intensive Care Med* 2016;31:319–24.
3 Bellomo R, Kellum JA, Ronco C. Acute kidney injury. *Lancet* 2012;380:756–66.
4 Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* 2012;120:c179–84.
5 Coelho S, Cabral G, Lopes JA, *et al*. Renal regeneration after acute kidney injury. *Nephrology* 2018;23:805–14.
6 Wilson FP, Greenberg JH. Acute kidney injury in real time: prediction, alerts, and clinical decision support. *Nephron* 2018;140:116–9.
7 Wu Y, Hao W, Chen Y, *et al*. Clinical features, risk factors, and clinical burden of acute kidney injury in older adults. *Ren Fail* 2020;42:1127–34.
8 Koyner JL, Adhikari R, Edelson DP, *et al*. Development of a multicenter ward-based AKI prediction model. *Clin J Am Soc Nephrol* 2016;11:1935–43.
9 Low S, Vathsala A, Murali TM, *et al*. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC Nephrol* 2019;20:32.
10 Tomašev N, Glorot X, Rae JW, *et al*. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019;572:116–9.
11 Flechet M, Falini S, Bonetti C, *et al*. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care* 2019;23:282.
12 Churpek MM, Carey KA, Edelson DP, *et al*. Internal and external validation of a machine learning risk score for acute kidney injury. *JAMA Netw Open* 2020;3:e2012892.
13 Yuan Q, Zhang H, Deng T, *et al*. Role of artificial intelligence in kidney disease. *Int J Med Sci* 2020;17:970–84.

14 Li Y, Xu J, Wang Y, *et al*. A novel machine learning algorithm, Bayesian networks model, to predict the high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol* 2020;43:752–61.

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;26:563–73.

16 James MT, Pannu N, Hemmelgarn BR, *et al*. Derivation and external validation of prediction models for advanced chronic kidney disease following acute kidney injury. *JAMA* 2017;318:1787–97.

17 Johnson AEW, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.

18 Thoral PJ, Peppink JM, Driessen RH, *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;49:e563–77.

19 Koyner JL, Carey KA, Edelson DP, *et al*. The development of a machine learning inpatient acute kidney injury prediction model. *Crit Care Med* 2018;46:1070–7.

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.
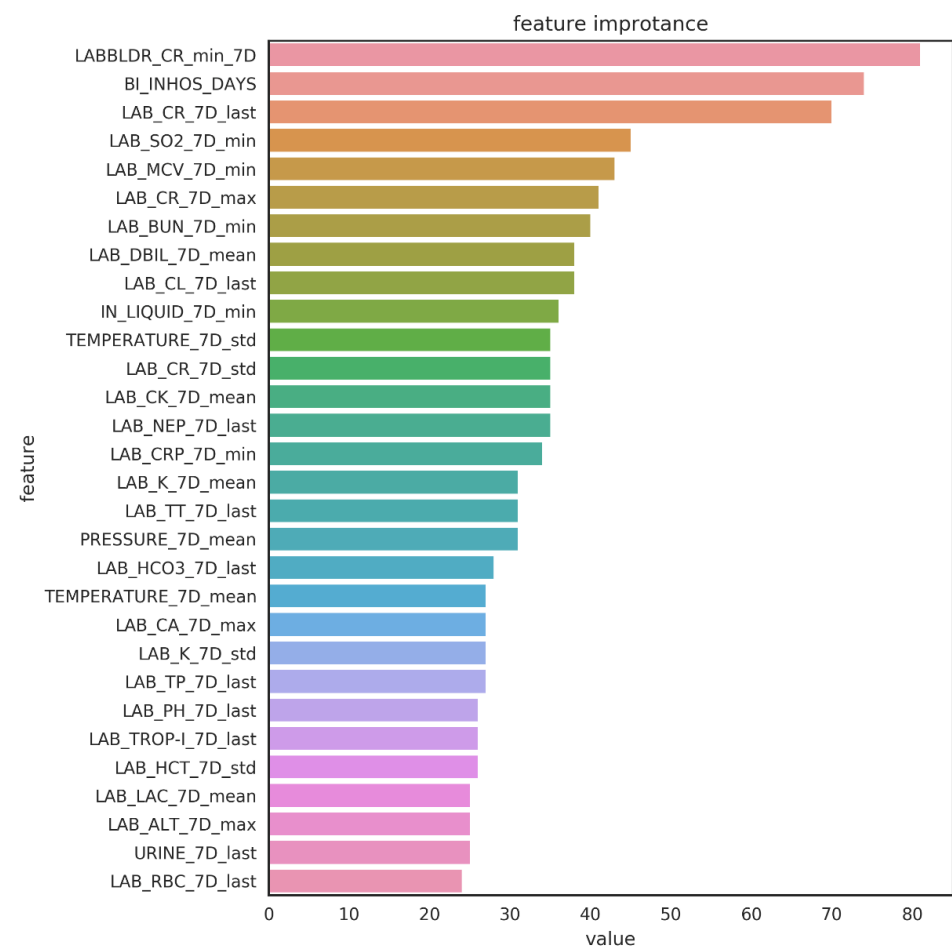
File name: **Figure S2.tif**

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.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

BMJ Open

KS Statistic Plot
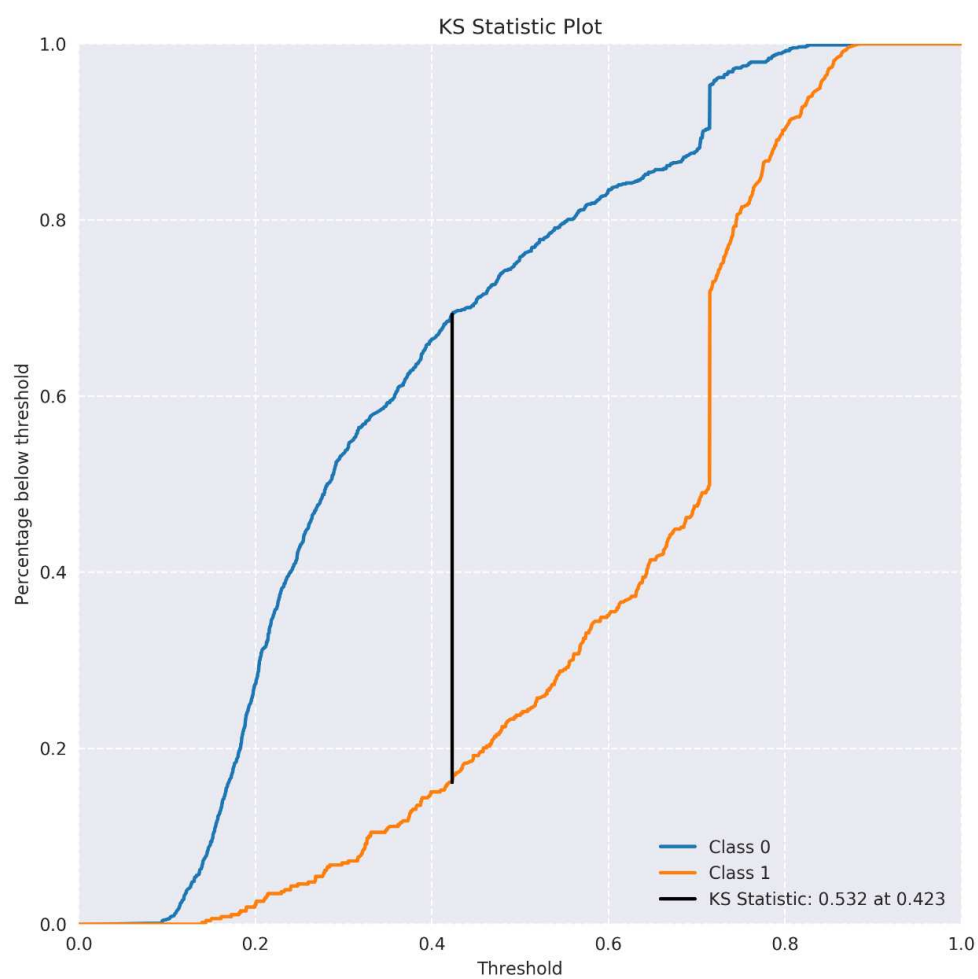
File name: **Table S1.docx**

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	√	√	√	√
PaCO₂	√	√	√	√
PaO₂	√	√	√	√

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.