Machine learning for the prediction of acute kidney injury in patients with acute pancreatitis admitted to the intensive care unit

Yisong Cheng[1], Jie Yang[1], Qin Wu[1], Lili Cao[2], Bo Wang[1], Xiaodong Jin[1], Yan Kang[1], Zhongwei Zhang[1], Min He[1]

[1]Department of Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China;

[2]Chengdu University, Institute of Basic Medical College, Chengdu, Sichuan 610106, China.

To the Editor: Acute kidney injury (AKI) is a well-known and critical complication of acute pancreatitis in intensive care units (ICUs). Patients with acute pancreatitis who developed AKI had a significantly higher mortality than those without AKI.[1] Currently, the diagnosis of AKI in acute pancreatitis patients includes a sudden decrease in glomerular filtration rate manifested by an increase in serum creatinine or oliguria within 48 h to 7 days.[2] However, kidney damage occurs unstoppably when creatinine increases or urine output decreases. This study aimed to establish prediction models based on the machine learning algorithm and traditional logistic regression method only with commonly collected variables when admitted to the ICU, and we also compared their performance for predicting the occurrence of AKI in patients with acute pancreatitis.

This study used a retrospective database of patients with acute pancreatitis admitted to the ICU in West China Hospital, Sichuan University from December 2015 to December 2019. The study complied with the *Declaration of Helsinki*, and the Institutional Review Board of West China Hospital approved the protocol. The Ethics Committee exempted informed consent since it was a retrospective study. The exclusion criteria were as follows: (1) age <18 years; (2) already diagnosed with AKI when admitted to the ICU; (3) history of chronic kidney disease; and (4) pancreas condition (pancreatic trauma, tumor, and chronic pancreatitis).

AKI was diagnosed using serum creatinine and urine output according to the definition of Kidney disease: Improving global outcomes (KDIGO) guidelines.[3] Acute pancreatitis was defined as at least two or more of the following three features: (1) typical acute pancreatitis abdominal pain; (2) elevated lipase (amylase or lipase) at least three times the upper limit of normal; and (3) imaging manifestations.[4]

We divided the samples into a training set and testing set. Approximately 70% of patients were used for model training, and the remaining 30% were used for model validation to avoid repeated analysis of the same sample. Furthermore, a 10-fold cross-validation method was also adopted to balance the computational time and variance. In the training process, the optimal hyperparameters were chosen with a grid search method of optimizing hypermeters through exhaustive search.

We selected logistic regression to build the traditional model. First, univariate logistic regression was used to identify potential risk factors, and then eight variables with $P < 0.05$ were simultaneously entered into the multiple logistics regression (LR) model (forward stepwise logistic regression), in which the importance was estimated by the standard correlation coefficient. Parameters with $P < 0.05$ were selected to construct the LR model, and a prediction nomogram was plotted based on the multiple LR analysis.

Machine learning derives a mathematical model with a non-ruled-based approach. First, the machine learning models were supplied in the training set, and some model parameters were chosen arbitrarily. In the training steps, the model gradual tunes trainable parameters to optimize the performance by itself. We applied gradient boosting (GB), random forest (RF), and extreme gradient boosting (XGB) since these common ensemble models have relatively accurate prediction performance, and more importantly, they are explainable.

A total of 488 patients diagnosed with acute pancreatitis at ICU admission were enrolled in this research and 151 (30.9%) patients developed AKI. In the traning set, there were 108 patients with AKI and 233 patients without AKI. On average, acute pancreatitis patients with AKI had higher levels of white blood cells, neutrophils, creatinine, uric acid, cystatin C, amylase, lipase, activated partial

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000002531

Yisong Cheng and Jie Yang contributed equally to this work.

Correspondence to: Prof. Zhongwei Zhang, Department of Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China
E-Mail: zhangzhongwei@scu.edu.cn;
Prof. Min He, Department of Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China
E-Mail: hemin19910306@wchscu.cn

Chinese Medical Journal 2022;135(23)

Received: 19-05-2022; **Online:** 03-01-2023 **Edited by:** Yanjie Yin

thromboplastin time, D-dimer, procalcitonin, C-reactive protein, ICU length, and 28-day mortality than those who did not have AKI [Supplementary Table 1, http://links.lww.com/CM9/B385].

Univariate logistic regression analysis revealed that eight variables were associated with the occurrence of AKI. After adjusting for potential risk factors in multivariable regression analysis, procalcitonin (odds ratio [OR] = 1.024, 95% confidence interval [CI]: 1.007–1.040, $P = 0.004$), natural logarithm B-type natriuretic peptide (Ln BNP) (OR = 1.010, 95% CI: 1.005–1.031, $P = 0.013$), and creatinine (OR = 1.004, 95% CI: 1.003–1.006, $P < 0.001$) were independently associated with AKI [Supplementary Table 2, http://links.lww.com/CM9/B385]. We then constructed a nomogram for the prediction of AKI according to the variables screened. The value of each variable corresponds to a score on the scale axis, and the total score (0–160) was determined based on the individual scores calculated using the nomogram [Supplementary Figure 1, http://links.lww.com/CM9/B385].

The area under the curve (AUCs) of LR, GB, RF, and XGB in predicting AKI were 0.763, 0.828, 0.812, and 0.809, respectively, $P < 0.001$ [Supplementary Figure 2, http://links.lww.com/CM9/B385]. Compared with the LR model, all three machine learning models had a higher AUC, but the difference was not statistically significant ($P > 0.05$) [Supplementary Table 3, http://links.lww.com/CM9/B385]. We further evaluated the potential improved discrimination and reclassification using the category-based net reclassification improvement (NRI) and found that all machine learning models had a significant NRI and integrated discrimination improvement (IDI) discrimination improvement ($P < 0.05$).

Likewise, the decision curve analysis [Figure 1] demonstrated that the net benefit of machine learning models surpassed that of the LR model when the threshold probability was approximately less than 0.4, indicating that machine learning-based prediction would more accurately identify the occurrence of AKI in patients with acute pancreatitis.

We summarized the top 20 important features of the GB model. According to the prediction model, a higher SHAP value indicates an increased risk of AKI development. Creatinine, uric acid, procalcitonin, thrombin time, and BNP were the most important predictors for AKI in the GB model [Supplementary Figure 3, http://links.lww.com/CM9/B385].

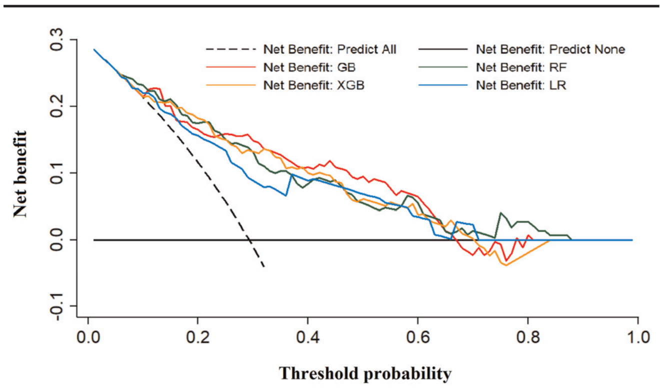

Figure 1: Decision curve analysis for machine learning and logistic regression model. GB: Gradient boosting; LR: Logistic regression; RF: Random forest; XGB: Extreme gradient boosting.

In the current study, we applied both logistic regression and machine learning models for the prediction of AKI in patients with acute pancreatitis. Compared to the traditional modeling approach, the machine learning method demonstrated superior performance, including net reclassification and net benefit, although the improvement in AUC was not statistically significant.

The current study has several limitations. First, our analysis used a relatively small retrospective dataset from a single center to train the models and lacked external validation for its generalizability. Second, some helpful clinical variables (eg, urine output, comorbidity, chronic medication, and treatment information) are lacking because of unavailable data. Third, the prediction only used the data when admitted to the ICU, and the diagnosing date of AKI was lacking, thus it is not known how far in advance it was predicted.

In conclusion, this retrospective cohort study developed three machine learning models and a traditional logistic regression model for the prediction of AKI in patients with acute pancreatitis admitted to the ICU. We found that machine learning was a superior conventional approach. The machine learning model may become a useful tool to support clinical decisions if the model is further validated.

Source of funding

The work was supported by grants from the National Natural Science Foundation of China (Nos. 81801892, 81901998) and the 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (No. ZYGD18020).

Conflicts of interest

None.

References

1. Shi N, Sun GD, Ji YY, Wang Y, Zhu YC, Xie WQ, *et al*. Effects of acute kidney injury on acute pancreatitis patients' survival rate in intensive care unit: A retrospective study. World J Gastroenterol 2021;27:6453–6464. doi: 10.3748/wjg.v27.i38.6453.
2. Levey AS, Eckardt KU, Dorman NM, Christiansen SL, Hoorn EJ, Ingelfinger JR, *et al*. Nomenclature for kidney function and disease: report of a kidney disease: Improving global outcomes (KDIGO) consensus conference. Kidney Int 2020;97:1117–1129. doi: 10.1016/j.kint.2020.02.010.
3. Kellum JA, Lameire N, Aspelin P, Barsoum RS, Burdmann EA, Goldstein SL, *et al*. Kidney disease: Improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. Kidney Int 2012;2:1–138.
4. Crockett SD, Wani S, Gardner TB, Falck-Ytter Y, Barkun AN. American gastroenterological association institute guideline on initial management of acute pancreatitis. Gastroenterology 2018;154:1096–1101. doi: 10.1053/j.gastro.2018.01.032.

How to cite this article: Cheng Y, Yang J, Wu Q, Cao L, Wang B, Jin X, Kang Y, Zhang Z, He M. Machine learning for the prediction of acute kidney injury in patients with acute pancreatitis admitted to the intensive care unit. Chin Med J 2022;135:2886–2887. doi: 10.1097/CM9.0000000000002531