



# A new prediction model for acute kidney injury in patients with sepsis

Chenyu Fan<sup>^</sup>, Xiu Ding, Yanli Song

Department of Emergency, Tongji Hospital of Tongji University, Shanghai, China

*Contributions:* (I) Conception and design: Y Song, C Fan; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: X Ding; (V) Data analysis and interpretation: C Fan, X Ding; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Yanli Song, MD. Department of Emergency, Tongji Hospital of Tongji University, 389#Xincun Road, Shanghai 200065, China. Email: songyanli@tongji.edu.cn.

**Background:** Acute kidney injury is common in patients with sepsis and contributes to poor prognosis and mortality. Early identification of high-risk patients can provide evidence for clinical decision-making.

**Methods:** We developed a prediction model based on a cohort of 15,726 patients with sepsis from the Medical Information Mart for Intensive Care III critical care database. Logistic regression analysis was applied to develop a prediction model incorporating the selected risk factors. Discrimination and calibration of the prediction model were assessed using the C-index and calibration plot.

**Results:** Risk factors in the prediction model included diabetes mellitus, chronic kidney disease, congestive heart failure, chronic liver disease, hyperbicarbonemia, hyperglycemia, low blood pH, prolonged clotting time, hypotension, and hyperlactatemia. The model showed great discrimination with a C-index of 0.711 (95% CI, 0.702–0.721) and appropriate calibration. A medium C-index value of 0.712 (95% CI, 0.697–0.727) could still be reached in the validation cohort. Negative and positive predictive values for the optimal cutoff value of  $\geq 6$  points were 56.8% and 72.3% in the training cohort and 57.3% and 72.6% in the validation cohort, respectively.

**Conclusions:** This prediction model allows clinicians to quickly assess the risk of sepsis-associated acute kidney injury (SA-AKI) at an early stage. Accordingly, clinicians can implement more medical measures that are considered beneficial to patients with sepsis.

**Keywords:** Sepsis; risk factors; acute kidney injury; prediction model

Submitted May 14, 2020. Accepted for publication Oct 23, 2020.

doi: 10.21037/apm-20-1117

View this article at: <http://dx.doi.org/10.21037/apm-20-1117>

## Introduction

Sepsis is a severe disease characterized by multiple organ injuries due to the patient's excess response to infection (1). The kidneys are the most common organ injured by sepsis, resulting in sepsis-associated acute kidney injury (SA-AKI). SA-AKI contributes to high mortality and poor prognosis in patients with sepsis (2), and is clinically nonspecific. Therefore, quick identification of patients with AKI is

difficult in the treatment of sepsis. In some studies, new biomarkers of kidney injury, such as urinary neutrophil gelatinase-associated lipocalin and urinary exosomal activating transcriptional factor 3 (3,4), were used to identify high-risk patients. However, these biomarkers have insufficient clinical application.

Clinical prediction models estimate the probability of risk by statistically combining a set of characteristics of the

<sup>^</sup> ORCID: 0000-0002-5405-3090.

patient and the disease. It could be a useful clinical support tool, assisting clinicians in complicated decision-making of medical treatments. Several clinical prediction models have been developed to predict AKI in specific clinical scenarios (5-13). Palomba *et al.* (5-7) established some predictive models to predict AKI risk after cardiac surgery, and Kheterpal *et al.* (8-9,12) were more concerned about the outcome after general surgery. Some studies (10-11) have described the incidence of AKI in intensive care unit (ICU) patients. However, the predictive power of these models was limited because of small sample size and the choice of a certain population. Models evaluating the clinical risk for the development of AKI in patients with sepsis are insufficient. We aimed to develop a new clinical tool to assess the risk of SA-AKI and provide more information for prevention, early diagnosis, and targeted interventions. We present the following article in accordance with the STROBE reporting checklist (available at <http://dx.doi.org/10.21037/apm-20-1117>).

## Methods

### *Study design and data source*

In this prediction model development and validation study, we registered and used the Medical Information Mart for Intensive Care III database (MIMIC-III). MIMIC-III is an extensive, freely available database comprising health-related data associated with over 40,000 patients who were admitted in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012 (14). MIMIC-III offers patients' private information, which is made available to researchers worldwide. We gained access to the database after completing the training courses at the National Institutes of Health. The study was approved by Massachusetts Institute of Technology Affiliates. (NO. 27653720). Requirements for written informed consent were waived because all personal data in this database were de-identified before the analyses. The study conforms to the provisions of the Declaration of Helsinki (as revised in 2013).

We identified patients with sepsis according to the sepsis-3 diagnostic criteria (1). We used the International Classification of Diseases diagnosis codes to identify infected patients, and infected patients with a Sequential Organ Failure Assessment (SOFA) score  $\geq 2$  were diagnosed with sepsis (15). A cohort of 15,726 adult patients with sepsis was selected for our study.

### *Variable selection*

Studies comprising patients in general wards diagnosed with AKI have led to our understanding of SA-AKI. We suspected that the risk factors for AKI identified in a general population are likely to confer an equal or even higher risk in patients with sepsis (2). Based on the results of published studies (5-10), we chose the following predictor variables that are assessable at early hospital admission: (I) Chronic comorbidities including advanced age ( $>70$  years), diabetes mellitus, congestive heart failure, chronic kidney disease, and chronic liver disease and (II) Acute events included hypotension (mean arterial pressure  $<70$  mmHg or use of any vasopressor), hyperglycemia (blood glucose  $>120$  mg/dL), anemia (hemoglobin  $<9$  mg/dL), decreased albumin (serum albumin  $<3$  mg/dL), hyperbicarbonemia (bicarbonate  $>30$  mg/dL), fever ( $T >37.3$  °C), low blood pH ( $<7.30$ ), hypoxemia (pressure of oxygen  $<60$  mmHg or use of mechanical ventilation), elevated bilirubin (serum bilirubin  $>2$  mg/dL), hyperlactatemia (blood lactate  $>2.1$  mmol/L), elevated white blood cells (white blood cells  $>10 \times 10^9/L$ ), decreased platelet (platelet count  $<150$  K/uL), and prolonged clotting time (partial thromboplastin time  $>35$  s).

Considering that it is challenging to interpret the risk degree of the continuous variables in a logistic model (16), all continuous variables were converted into categorical variables according to the standard clinical cutoff points. In total, 18 binary predictor variables were selected for model development.

### *Outcome measure*

The outcome was the diagnosis of AKI within 48 hours after ICU admission. In our study, we used clinical practice guidelines presented by the Kidney Disease: Improving Global Outcomes for AKI as the diagnostic criteria (17). The MIMIC-III database only comprises data after ICU admission, and the baseline data of patients' serum creatinine are unknown. Therefore, the first measurement after admission was taken as the baseline value, and when the serum creatinine levels of patients increased to  $>1.5$  times the baseline value within 48 hours, the patients were diagnosed with AKI. We did include urine volume in the diagnostic data because of missing data and difficulty in analysis.

### *Model development and validation*

Randomization was created using R software and was

divided into the training and validation cohort with a 7:3 allocation using random sequence. The least absolute shrinkage and selection operator (LASSO) method was used to select the optimal predictive features in risk factors from patients in the training cohort (18). Subsequently, a multivariate logistic regression analysis was used to develop a prediction model by incorporating the risk factors selected in the LASSO model. In the final multivariate model, the coefficients generated for each variable were approximated to the closest value to develop a simple SA-AKI risk score.

By summing all variables together, the total score can range from a minimum of 0 to a maximum of 16 points. The simplified score model was then assessed in the two cohorts. The optimal cutoff point for the risk score was determined by the C-index calculated in the training cohort. Discrimination was calculated using the C-index, ranging from 0.5 to 1.0 (1.0 means perfect discrimination). Calibration was usually described using calibration curves. We can observe that appropriate calibration is implied by a 45° diagonal line, whereas the relevant deviation above or below this line reflects underprediction or overprediction. Discrimination and calibration were assessed in the two cohorts.

### Statistical analysis

Data were missing for all risk factors except for age, sex, and chronic disease records. We filled in missing data using the technique of multiple imputations (19); accordingly, samples values were imputed based on posterior predictive distributions of missing data. We assessed that data were missing at random using a pattern graph. All analyses were performed using R software (version 3.6.3) and PostgreSQL software (version 10.0).  $P < 0.05$  was considered statistically significant in analyses.

## Results

### Patients' characteristics

A total of 15,726 patients with sepsis (218 patients excluded because of age) were admitted in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. According to the random principle, they were divided into the training cohort ( $n=11,008$ ) and the validation cohort ( $n=4,718$ ). All patients' data, including clinical and demographic characteristics, in the two sets are summarized

in *Table 1*.

### Model development

Of the chronic comorbidities and acute events, 19 risk factors were reduced to 10 potential predictors in the training cohort with nonzero coefficients in the LASSO regression model (when lambda was 0.01567499). These risk factors included diabetes mellitus, chronic kidney disease, congestive heart failure, chronic liver disease, hyperbicarbonemia, hyperglycemia, low blood pH, prolonged clotting time, hypotension, and hyperlactatemia.

For the collinearity diagnostics of these 10 variables, the VIF values were calculated to be  $< 5$ , which proved that there was no apparent collinearity between them. The results of the logistic regression analysis among the 10 risk factors are presented in *Table 2*.

The results of the final model can be described by the following equation: probability of SA-AKI =  $e^a / (1 + e^a)$ , where  $a = (\text{diabetes mellitus} \times 0.248) + (\text{chronic kidney disease} \times 0.994) + (\text{congestive heart failure} \times 0.069) + (\text{chronic liver disease} \times 0.181) + (\text{hyperbicarbonemia} \times 0.665) + (\text{hyperglycemia} \times 0.281) + (\text{low blood pH} \times 0.378) + (\text{prolonged clotting time} \times 0.231) + (\text{hypotension} \times 0.236) + [\text{hyperlactatemia} \times 0.292] - 1.039$ . The C-index of the model in the training cohort was 0.711 (95% CI, 0.702–0.721), and the sensitivity and specificity were 74.4% and 53.4%, respectively.

For more convenient clinical use, we converted the coefficients in the original model into risk scores, as described in the Methods section and shown in *Table 3*. There was no significant difference in the C-index 0.712 (95% CI, 0.697–0.727) after converting the regression coefficient-based model to the risk score. Two models showed appropriate calibration in the training cohort, with a reasonable agreement between observation and prediction.

### Internal validation

The original and simplified score models were validated in the validation cohort. The original model demonstrated good accuracy in estimating the risk of SA-AKI, with a C-index of 0.712 (95% CI, 0.697–0.727). Furthermore, the simplified score model showed a C-index of 0.705 (95% CI, 0.690–0.720) in the cohort (*Figure 1*). The calibration curve of the two models also demonstrated good agreement in the validation cohort (*Figure 2*).

A cutoff of  $\geq 6$  points was chosen based on the best

**Table 1** Clinical and demographic characteristics in the two sets

Risk factors	Training cohort (n=11,008)	Validation cohort (n=4,718)	Total (n=15,726)
AKI, n [%]	6,683 [61]	2,882 [61]	9,565 [61]
Advanced age, n [%]	5,383 [49]	2,361 [50]	7,744 [49]
Gender [male], n [%]	5,934 [54]	2,529 [54]	8,463 [54]
Diabetes mellitus, n [%]	3,793 [34]	1,684 [36]	5,477 [35]
Congestive heart failure, n [%]	6,161 [56]	2,666 [57]	8,827 [56]
Chronic kidney disease, n [%]	4,655 [42]	2,020 [43]	6,675 [42]
Chronic liver disease, n [%]	1,553 [14]	664 [14]	2,217 [14]
Hypotension, n [%]	3,703 [34]	1,541 [33]	5,244 [33]
Hyperglycemia, n [%]	3,219 [29]	1,399 [30]	4,618 [29]
Anemia, n [%]	4,411 [40]	1,876 [40]	6,287 [40]
Decreased albumin, n [%]	6,418 [58]	2,832 [60]	9,250 [59]
Hyperbicarbonemia, n [%]	6,440 [59]	2,787 [60]	9,227 [59]
Fever, n [%]	2,680 [24]	1,126 [24]	3,806 [24]
Low blood pH, n [%]	4,569 [42]	2,020 [43]	6,589 [42]
Hypoxemia, n [%]	2,308 [21]	950 [20]	3,258 [21]
Elevated bilirubin, n [%]	2,324 [21]	1,004 [21]	3,328 [21]
Hyperlactatemia, n [%]	5,891 [54]	2,516 [53]	8,407 [53]
Elevated white blood cells, n [%]	8,049 [73]	3,441 [73]	11,490 [73]
Decreased platelet, n [%]	4,319 [39]	1,843 [39]	6,162 [39]
Prolonged clotting time, n [%]	5,356 [49]	2,324 [49]	7,680 [49]

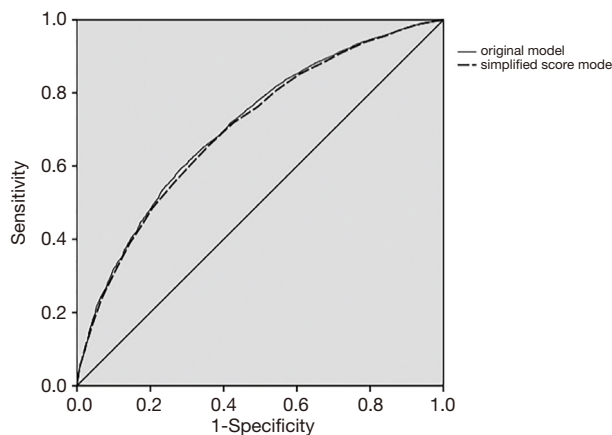
**Table 2** Predictors of AKI obtained by multivariate logistic regression analysis in the training cohort

Risk factors	Coefficient	Odds ratio	97.5% CI for OR		P value
			Lower	Upper	
Diabetes mellitus	0.248	1.282	1.170	1.405	<0.001
Chronic kidney disease	0.994	2.702	2.475	2.952	<0.001
Congestive heart failure	0.069	1.071	0.985	1.165	0.109
Chronic liver disease	0.181	1.198	1.057	1.360	0.005
Hyperbicarbonemia	0.665	1.945	1.785	2.119	<0.001
Hyperglycemia	0.281	1.324	1.200	1.461	<0.001
Low blood pH	0.378	1.459	1.337	1.590	<0.001
Prolonged clotting time	0.231	1.260	1.157	1.371	<0.001
Hypotension	0.236	1.267	1.159	1.383	<0.001
Hyperlactatemia	0.292	1.339	1.228	1.459	<0.001

**Table 3** SA-AKI risk prediction score of the final model

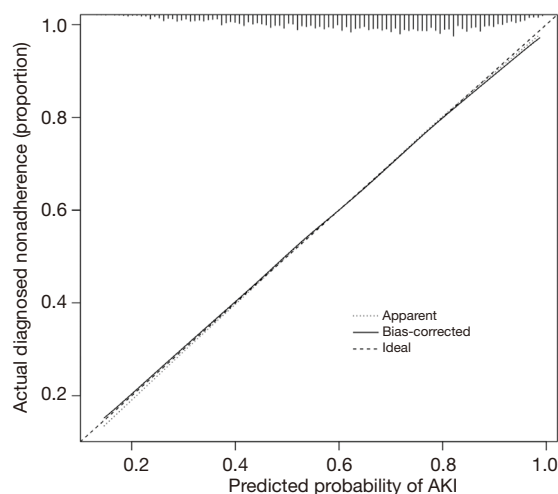
Risk factors	Points
Chronic	
Diabetes mellitus	1.5
Chronic kidney disease	3
Congestive heart failure	1
Chronic liver disease	1
Acute	
Hyperbicarbonemia	2
Hyperglycemia	1.5
Low blood pH	1.5
Prolonged clotting time	1.5
Hypotension	1.5
Hyperlactatemia	1.5

Minimum total score, 0; maximum total score, 16. SA-AKI, sepsis-associated acute kidney injury.



**Figure 1** The receiver operating characteristic curve in the validation cohort. The discriminative ability of the two models for the risk prediction of sepsis-associated acute kidney injury was expressed as a C-index for the validation cohort. The one with the larger C-index, the original model, is drawn in solid line.

Youden Index by a combination of sensitivity and specificity to identify high-risk patients (20). Negative and positive predictive values for the optimal cutoff value of  $\geq 6$  points were 56.8% and 72.3% in the training cohort and 57.3% and 72.6% in the validation cohort, respectively.



**Figure 2** Calibration curves of the model in the validation cohort. The solid line represents the performance of the nomogram. The diagonal dotted line represents a perfect prediction using an ideal model.

## Discussion

We developed a simplified score model to predict the risk of AKI in patients with sepsis. The model was successfully internally validated and showed good discrimination and calibration. Moreover, the optimal cutoff for the diagnosis of SA-AKI was estimated to be 6 points. The model incorporates 10 factors: diabetes mellitus, chronic kidney disease, congestive heart failure, chronic liver disease, hyperbicarbonemia, hyperglycemia, low blood pH, prolonged clotting time, hypotension, and hyperlactatemia. We carefully selected a list of candidate risk factor; this process entailed exclusion of proven risk factors with missing data in MIMIC-III. Future studies should include more variables such as biomarkers of kidney injury (21).

This study has several strengths. To our knowledge, the study included the largest cohort of patients with sepsis. We identified patients with sepsis by calculating the SOFA scores according to the sepsis-3 diagnostic criteria instead of the International Classification of Diseases codes in the database (1). The good performance of the simplified score model in the validation cohort supports the use of this model in patients with sepsis. A simplified score model is a practical tool that can be used in the early stages, and the risk factors are easily measured and routinely available. In addition to verifying the reported risk factors, we also

report two new factors [hyperlactatemia (22) and prolonged clotting time] affecting the development of SA-AKI. However, more well-designed randomized controlled trials are needed to further confirm these two risk factors.

The model is only applicable to patients with sepsis because sepsis is a significant risk factor in other AKI models, which often leads us to underestimate the role of other factors (10). Besides, in contrast to other AKI models (5,7,10) with a high negative predictive value, our model has a higher positive predictive value. This means that we can identify high-risk patients earlier.

The simplified score model has several practical applications. First, it is a small step toward more personalized medicine because it could identify individuals with sepsis at high-risk of AKI. However, at present, no more aggressive treatments are available for these patients (23). We believe that it is necessary to pay more attention to high-risk patients and that the nephrologists should carefully evaluate the timing of renal replacement therapy because some single-center retrospective studies have shown that patients can benefit from early renal replacement therapy (24). Second, patients and their relatives need useful information about the risk of SA-AKI so that they can participate efficiently in decision-making. They are less likely to depend only on the clinician's intuition in making decisions. Our simple score model can provide data to support these decisions. Third, the score model can be constructed into the hospital information system, and when the clinical information is updated, the patient's score will be reassessed automatically.

This study has several limitations. First, missing data were handled with multiple imputation techniques, which may lead to a decrease in the accuracy of the final model and a decrease in the C-index. Second, the urine criterion was not applied to diagnose AKI in our study because the data of urine per hour were not readily available, and this may decrease the overall incidence of AKI (25). Third, baseline serum creatinine values were difficult to obtain; thus, patient's condition before admission could not be appropriately determined. Fourth, our model incorporates 10 factors, the collection of these factors is time-consuming, and some of the indicators need to be measured by blood sampling. However, data for these indicators can be easily obtained in the emergency room and will not be a barrier to implementing the model. Finally, the lack of external validation of this model makes it difficult for broad applicability.

## Conclusions

By combining 10 risk factors of SA-AKI, a simplified score model was constructed for risk estimation. By estimating the risk of individual SA-AKI early, clinicians can implement more measures that are considered beneficial to patients with sepsis. Future studies are needed to externally validate our model using a new cohort of patients with sepsis.

## Acknowledgments

*Funding:* This study was funded by the Emergency and Critical Illness Important Weak Discipline Project of Shanghai Municipal Health Commission (2016ZB0204).

## Footnote

*Reporting Checklist:* The authors have completed the STROBE reporting checklist. Available at <http://dx.doi.org/10.21037/apm-20-1117>

*Data Sharing Statement:* Available at <http://dx.doi.org/10.21037/apm-20-1117>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/apm-20-1117>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by Massachusetts Institute of Technology Affiliates. (NO. 27653720). Requirements for written informed consent were waived because all personal data in this database were de-identified before the analyses. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013).

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the

formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

- Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801-10.
- Poston JT, Koyner JL. Sepsis associated acute kidney injury. *BMJ* 2019;364:k4891.
- Mussap M, Noto A, Fravega M, et al. Urine neutrophil gelatinase-associated lipocalin (uNGAL) and netrin-1: are they effectively improving the clinical management of sepsis-induced acute kidney injury (AKI)? *J Matern Fetal Neonatal Med* 2011;24 Suppl 2:15-7.
- Panich T, Chanchaoentana W, Somparn P, et al. Urinary exosomal activating transcriptional factor 3 as the early diagnostic biomarker for sepsis-induced acute kidney injury. *BMC Nephrol* 2017;18:10.
- Palomba H, De Castro I, Neto AL, et al. Acute kidney injury prediction following elective cardiac surgery: AKICS Score. *Kidney Int* 2007;72:624-31.
- Thakar CV, Arrigain S, Worley S, et al. A clinical score to predict acute renal failure after cardiac surgery. *J Am Soc Nephrol* 2005;16:162-8.
- Kiers HD, Van Den Boogaard M, Schoenmakers MC, et al. Comparison and clinical suitability of eight prediction models for cardiac surgery-related acute kidney injury. *Nephrol Dial Transplant* 2013;28:345-51.
- Kheterpal S, Tremper KK, Heung M, et al. Development and validation of an acute kidney injury risk index for patients undergoing general surgery: results from a national data set. *Anesthesiology* 2009;110:505-15.
- Kheterpal S, Tremper KK, Englesbe MJ, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. *Anesthesiology* 2007;107:892-902.
- Malhotra R, Kashani KB, Macedo E, et al. A risk prediction score for acute kidney injury in the intensive care unit. *Nephrol Dial Transplant* 2017;32:814-22.
- Kim WH, Park MH, Kim HJ, et al. Potentially modifiable risk factors for acute kidney injury after surgery on the thoracic aorta: a propensity score matched case-control study. *Medicine (Baltimore)* 2015;94:e273.
- Cho E, Kim SC, Kim MG, et al. The incidence and risk factors of acute kidney injury after hepatobiliary surgery: a prospective observational study. *BMC Nephrol* 2014;15:169.
- Chawla LS, Abell L, Mazhari R, et al. Identifying critically ill patients at high risk for developing acute renal failure: A pilot study. *Kidney Int* 2005;68:2274-80.
- Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
- Angus DC, Linde-Zwirble WT, Lidicker J, et al. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit Care Med* 2001;29:1303-10.
- Barrio I, Arostegui I, Quintana JM, et al. Use of generalised additive models to categorise continuous variables in clinical prediction. *BMC Med Res Methodol* 2013;13:83.
- Palevsky PM, Liu KD, Brophy PD, et al. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for acute kidney injury. *Am J Kidney Dis* 2013;61:649-72.
- Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* 2007;26:5512-28.
- Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol* 2010;171:624-32.
- Hilden J, Glasziou P. Regret graphs, diagnostic uncertainty and Youden's Index. *Stat Med* 1996;15:969-86.
- Moriates C, Maisel A. The utility of biomarkers in sorting out the complex patient. *Am J Med* 2010;123:393-9.
- Green JP, Berger T, Garg N, et al. Hyperlactatemia affects the association of hyperglycemia with mortality in nondiabetic adults with sepsis. *Acad Emerg Med* 2012;19:1268-75.
- Quenot JP, Dargent A, Large A, et al. Treatment of sepsis-induced acute kidney injury in the ICU: the therapeutic targets do not seem to be established yet. *Ann Transl Med* 2019;7:S181.
- Karvellas CJ, Farhat MR, Sajjad I, et al. A comparison of early versus late initiation of renal replacement therapy in critically ill patients with acute kidney injury: a systematic review and meta-analysis. *Crit Care* 2011;15:R72.
- Kellum JA, Sileanu FE, Murugan R, et al. Classifying AKI by Urine Output versus Serum Creatinine Level. *J Am Soc Nephrol* 2015;26:2231-8.

**Cite this article as:** Fan C, Ding X, Song Y. A new prediction model for acute kidney injury in patients with sepsis. *Ann Palliat Med* 2021;10(2):1772-1778. doi: 10.21037/apm-20-1117