

Original Paper

# Real-Time Prediction of Sepsis in Critical Trauma Patients: Machine Learning–Based Modeling Study

Jiang Li<sup>1\*</sup>, BSc; Fengchan Xi<sup>2,3\*</sup>, MD; Wenkui Yu<sup>3\*</sup>, MD; Chuanrui Sun<sup>1</sup>, BSc; Xiling Wang<sup>1</sup>, PhD

<sup>1</sup>School of Public Health and Key Laboratory of Public Health Safety, Fudan University, Shanghai, China

<sup>2</sup>Research Institute of General Surgery, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing, China

<sup>3</sup>Department of Intensive Care Unit, Affiliated Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China

\*these authors contributed equally

**Corresponding Author:**

Xiling Wang, PhD

School of Public Health and Key Laboratory of Public Health Safety

Fudan University

No 130 Dongan Road

Xuhui District

Shanghai, 200032

China

Phone: 86 021 54237051

Email: [erinwang@fudan.edu.cn](mailto:erinwang@fudan.edu.cn)

## Abstract

**Background:** Sepsis is a leading cause of death in patients with trauma, and the risk of mortality increases significantly for each hour of delay in treatment. A hypermetabolic baseline and explosive inflammatory immune response mask clinical signs and symptoms of sepsis in trauma patients, making early diagnosis of sepsis more challenging. Machine learning–based predictive modeling has shown great promise in evaluating and predicting sepsis risk in the general intensive care unit (ICU) setting, but there has been no sepsis prediction model specifically developed for trauma patients so far.

**Objective:** To develop a machine learning model to predict the risk of sepsis at an hourly scale among ICU-admitted trauma patients.

**Methods:** We extracted data from adult trauma patients admitted to the ICU at Beth Israel Deaconess Medical Center between 2008 and 2019. A total of 42 raw variables were collected, including demographics, vital signs, arterial blood gas, and laboratory tests. We further derived a total of 485 features, including measurement pattern features, scoring features, and time-series variables, from the raw variables by feature engineering. The data set was randomly split into 70% for model development with stratified 5-fold cross-validation, 15% for calibration, and 15% for testing. An Extreme Gradient Boosting (XGBoost) model was developed to predict the hourly risk of sepsis at prediction windows of 4, 6, 8, 12, and 24 hours. We evaluated model performance for discrimination and calibration both at time-step and outcome levels. Clinical applicability of the model was evaluated with varying levels of precision, and the potential clinical net benefit was assessed with decision curve analysis (DCA). A Shapley additive explanation algorithm was applied to show the effect of features on the prediction model. In addition, we trained an L2-regularized logistic regression model to compare its performance with XGBoost.

**Results:** We included 4603 trauma patients in the study, 1196 (26%) of whom developed sepsis. The XGBoost model achieved an area under the receiver operating characteristics curve (AUROC) ranging from 0.83 to 0.88 at the 4-to-24-hour prediction window in the test set. With a ratio of 9 false alerts for every true alert, it predicted 73% (386/529) of sepsis-positive timesteps and 91% (163/179) of sepsis events in the subsequent 6 hours. The DCA showed our model had a positive net benefit in the threshold probability range of 0 to 0.6. In comparison, the logistic regression model achieved lower performance, with AUROC ranging from 0.76 to 0.84 at the 4-to-24-hour prediction window.

**Conclusions:** The machine learning–based model had good discrimination and calibration performance for sepsis prediction in critical trauma patients. Using the model in clinical practice might help to identify patients at risk of sepsis in a time window that enables personalized intervention and early treatment.

(*JMIR Form Res* 2023;7:e42452) doi: [10.2196/42452](https://doi.org/10.2196/42452)

**KEYWORDS**

sepsis; trauma; intensive care unit; machine learning; real-time prediction

## Introduction

Sepsis is a life-threatening type of organ dysfunction caused by a dysregulated host response to an infection [1]. It is a major contributor to the global burden of disease, with morbidity and mortality rates having failed to decrease substantially during the past decade, especially in the trauma population [2,3]. According to the international consensus guidelines for sepsis, fluid resuscitation should commence within the first 3 hours of sepsis, and antimicrobial treatment should commence within 1 hour of sepsis [4-6]. The mortality rate of sepsis increases significantly with each hour of delayed administration of antibiotics [5,6]. However, early recognition of sepsis can be challenging due to the complexity of the sepsis response and the heterogeneity of the population with sepsis [7,8]. Furthermore, delays in communication among health care providers may exacerbate sepsis-management delays [9]. Therefore, closely evaluating and predicting the risk of sepsis before onset at an individual level may provide insights for clinicians to implement timely personalized medicine to improve prognoses.

The traditional tools to predict sepsis are often based on generalized linear models. The Epic Sepsis Model (ESM), a penalized logistic regression model, is one of the most widely implemented early warning systems for sepsis, especially in the United States. However, Wong et al [10] recently found that the ESM had poor discrimination performance, with an area under the receiver operating characteristics curve (AUROC) of 0.76 to predict sepsis 4 hours in advance; it also failed to detect sepsis before its onset in 67% of patients. Machine learning-based predictive modeling is increasingly popular and is being applied in clinical research and practice due to the availability of large digitized medical data sets and computing power [11,12]. The advantage of machine learning algorithms lies in their capability to extract the most important information from complex data and capture nonlinear relations between features. Machine learning models, including gradient boosting trees, random forests, and neural networks, have been developed for real-time prediction of sepsis or sepsis shock in a general intensive care unit (ICU) setting [13,14].

However, to our knowledge, there is no such real-time prediction model aimed specifically at the trauma population. Unlike general patients, most trauma patients are relatively young, are predominantly male, and have few underlying medical conditions [2,15]. The weight of these factors in the prediction models for trauma patients might differ from the weights in models developed for other critical patients. Furthermore, a hypermetabolic baseline and explosive inflammatory immune response mask clinical signs and symptoms of sepsis in trauma patients, making it more difficult to diagnose sepsis in the early stages [16,17]. Therefore, the development of a real-time prediction model for sepsis in the trauma population would be clinically valuable and could help clinicians to identify patients at high risk of developing sepsis, leading to improved medical care [13]. In this study, we aimed to develop a machine learning

model using Extreme Gradient Boosting (XGBoost) and a publicly available database to predict the risk of sepsis at an hourly scale in trauma patients admitted to an ICU.

## Methods

### Data Source

Data were obtained from a publicly available database, the Medical Information Mart for Intensive Care IV (MIMIC IV; version 1.0), which continuously collected medical records from the ICU at Beth Israel Deaconess Medical Center (Boston, MA) between 2008 and 2019 [18].

### Patient Selection and Variable Extraction

All patients aged  $\geq 18$  years in the database who had a first-discharge diagnosis of trauma according to the ninth or tenth revisions of the International Classification of Diseases (ICD) codes (ICD-9: 800-848, 850-854, 860-887, 890-897, 900-904, 910-929, or 950-957; ICD-10: S00-S99) were included. In the case of multiple ICU admissions, we used only data from the first episode of ICU admission to avoid repeated measures of sepsis. Patients who developed sepsis before ICU admission were excluded. Medical records after the occurrence of sepsis were not used in the model development due to considerations of the clinical applicability of the model.

A total of 42 raw variables were chosen based on the previous literature and their clinical relevance. They were extracted based on an SQL search with Navicat Premium (version 15.0.21; PremiumSoft CyberTech Ltd) [13]. These features represented a mix of static and dynamic information. A full set of the variables is listed in [Multimedia Appendix 1](#).

### Ethical Approval

This database was approved by the Beth Israel Deaconess Medical Center (45682859) [19]. The need for informed consent was waived because of the completely anonymous nature of the data and the retrospective nature of the study. We completed the relevant courses to access the database and obtained a certificate (45682859).

### Outcomes

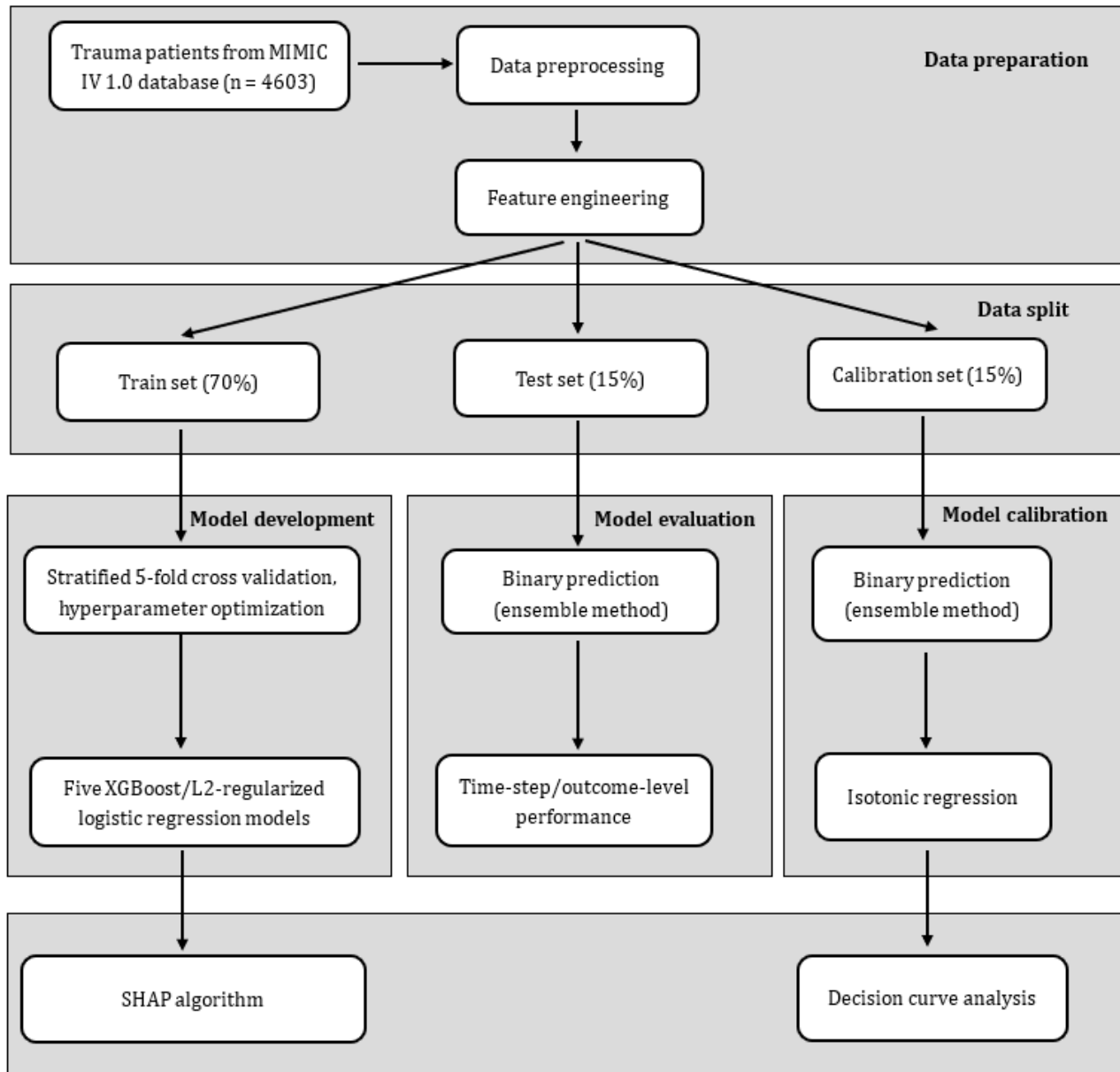
Sepsis was defined as the presence of both suspected infection and organ dysfunction according to the recent sepsis-3 criteria [1,20]. The onset time of sepsis was defined as the earliest time of suspected infection and organ dysfunction, manifested as an acute increase in the Sequential Organ Failure Assessment (SOFA) score of at least 2 [21,22]. More details on the definition of the onset time of sepsis are provided in [Multimedia Appendix 2](#).

### Data Preprocessing

To optimize the data for the model input, static variables were repeated at each 1-hour time grid. Dynamic variables measured more than once per hour were aggregated into 1-hour time steps by calculating hourly medians. We adopted the last-occurrence-carry-forward strategy to impute missing values

for each variable. Population means were used for imputing the remaining missing values occurring before the first measurement [23]. A schematic workflow of the study is shown in Figure 1.

**Figure 1.** Flowchart of model development. MIMIC IV: Medical Information Mart for Intensive Care IV; SHAP: Shapley additive explanation.



## Feature Engineering

A total of 485 features were derived from the raw variables, classified into three subtypes: (1) 37 measurement pattern features, (2) 7 scoring features, and (3) 441 time-series variables. Details of the feature engineering are described in [Multimedia Appendix 3](#). Finally, a total of 527 features were used for model development.

## Model Development

The data set was randomly split into three sub-data sets: 70% for model training with stratified 5-fold cross-validation, 15% for calibration, and 15% for testing. Records for each patient, rather than individual time steps, were assigned to the same training, validation, calibration, and test sets to avoid label leaking. As just over 2% (529/26,140) of individual time steps presented to the model were labeled as sepsis, we remedied this

imbalance in the data set by tuning parameters to change the weight between the positive and negative classes during the training process.

We used XGBoost, a gradient boosting algorithm well-known for obtaining winning solutions in various data competitions [24], to predict the risk of sepsis onset among trauma patients in the following prediction windows: 4, 6, 8, 12, and 24 hours; the temporal resolution was 1 hour. The choice of time windows was in accordance with previous literature predicting the risk of sepsis in general ICU patients [13,25] and takes into account the time needed before making interventions in clinical sepsis management, as well as prediction accuracy [26]. To reduce the risk of model overfitting, 5-fold cross-validation was used to produce 5 XGBoost models on the training set. Bayesian optimization was used to select the optimal hyperparameter combinations by maximizing AUROC in the validation set [27].

The ensemble method was used to provide robust estimation by averaging prediction probabilities from the above 5 models [28].

In addition, we trained an L2-regularized logistic regression model to compare its performance with XGBoost. Continuous features were standardized to improve the speed of model convergence before fitting. The grid search algorithm was applied to select the optimal strength of regularization. The ensemble approach was also adopted for the final prediction.

### Model Evaluation and Model Calibration

We evaluated model discrimination performance on the test set at both the time-step level and outcome level. At the time-step level, we calculated the AUROC and the area under the precision-recall curve (AUPRC) with prediction windows of 4, 6, 8, 12, and 24 hours for XGBoost and logistic regression. Sensitivity and specificity were calculated for prediction window/precision pairs (at 5%, 8%, and 10%) for XGBoost. At the outcome level, we computed sensitivity at different levels of precision. Unlike time-step-level sensitivity, outcome-level sensitivity corresponded to the percentage of all sepsis episodes that had at least one correct prediction within a specific time window before sepsis onset. Model calibration was evaluated with the average calibration error (ACE) [22] and reliability plots [29]. Isotonic regression was used to recalibrate the probability from the XGBoost model in the calibration set to obtain more accurate predictions [30]. Furthermore, a decision curve analysis (DCA) was conducted to assess the potential

benefit of guiding sepsis management based on predictions from our model across the threshold probabilities of 0 to 0.6. We set the upper limit of threshold probability at 0.6 because it is clinically unreasonable for a patient or doctor to accept a risk greater than 0.6 by balancing the harms of missing a patient with sepsis and unnecessary intervention on a patient without sepsis [31,32].

### Shapley Additive Explanation Algorithm

The Shapley additive explanation (SHAP) algorithm was used to show the average effect of each feature on the prediction model [33,34]. Bootstrapping was used to construct 95% CIs of the estimates using 1000 bootstrap samples of sepsis probabilities with replacement [23]. All computational analyses were conducted with Python (version 3.9.7; Python Software Foundation).

## Results

### Patient Characteristics

We obtained the medical records of 4603 trauma patients admitted to the ICU from MIMIC IV. After splitting the data randomly, there were 3222, 691, and 690 patients in the training, calibration, and testing sets, respectively. The 3 cohorts had similar characteristics, with a median age of 63 to 65 years and a higher proportion of males (ranging from 61% to 65%). The prevalence of sepsis in the above data sets was around 26% (Table 1).

**Table 1.** Characteristics of the trauma patients in the training, calibration, and testing sets.

Characteristics	All (N=4603)	Training set (n=3222)	Calibration set (n=691)	Testing set (n=690)
Age (years), median (IQR)	64 (42-81)	64 (42-81)	63 (44-81)	65 (42-82)
<b>Sex, n (%)</b>				
Male	2878 (62.5)	2015 (62.5)	418 (60.5)	445 (64.5)
Female	1725 (37.5)	1207 (37.5)	273 (39.5)	245 (35.5)
Charlson comorbidity index, median (IQR)	4 (1-5)	4 (1-5)	4 (1-5)	4 (1-5)
BMI (kg/m <sup>2</sup> ), median (IQR)	26.9 (26.9-26.9)	26.9 (26.9-26.9)	26.9 (26.9-26.9)	27.9 (27.6-27.9)
Time interval from hospital to ICU <sup>a</sup> admission (hours), median (IQR)	1 (0-1)	1 (0-1)	1 (0-1)	1 (0-2)
Length of stay in ICU (hours), median (IQR)	40 (21-82)	39 (21-82)	41 (21-84)	42 (21-81)
Sepsis, n (%)	1196 (26)	837 (26)	180 (26.1)	179 (25.9)

<sup>a</sup>ICU: intensive care unit.

### Model Evaluation and Model Calibration

In the test set, XGBoost outperformed logistic regression in both discrimination and calibration across all prediction windows (Table 2). For a prediction window of 6 hours, XGBoost had a higher AUROC (0.87, 95% CI 0.85-0.89), higher AUPRC (0.27, 95% CI 0.23-0.31) and lower ACE (0.33, 95% CI 0.31-0.35) than the logistic regression (AUROC=0.83, 95% CI 0.81-0.85; AUPRC=0.18, 95% CI 0.15-0.21; and ACE=0.44, 95% CI 0.44-0.45; Table 2, Multimedia Appendix 4). With longer prediction windows, the model discrimination as evaluated by AUROC or AUPRC decreased. The AUROC of

the XGBoost model decreased from 0.88 (95% CI 0.86-0.90) in the 4-hour prediction window to 0.83 (95% CI 0.81-0.84) in the 24-hour window, and the AUROC of the logistic regression model decreased from 0.84 (95% CI 0.82-0.86) to 0.76 (95% CI 0.74-0.77). However, the model calibration improved slightly with an increase in the prediction window. The ACE of the XGBoost model decreased from 0.35 (95% CI 0.32-0.37) in the 4-hour prediction window to 0.30 (95% CI 0.28-0.32) in the 24-hour window, and the ACE of the logistic regression model decreased from 0.45 (95% CI 0.44-0.45) to 0.42 (95% CI 0.42-0.43; Table 2).

At the time-step level, by using XGBoost, 73% (386/529) of sepsis-positive time steps were predicted at the 6-hour prediction window with a ratio of 9 false predictions for every true positive (10% precision), while 81% (428/529) of sepsis-positive time steps were predicted with a ratio of 12 false predictions for every true positive (8% precision; [Figure 2](#)). At the outcome level, the proportion of predicted sepsis episodes decreased with increased precision level. At the 10% precision level, XGBoost identified 91% (163/179) of sepsis events occurring in the subsequent 6 hours. Of note, the total number of events to be

identified became fewer as the time period became shorter. There was a total of 22% (40/179) of patients for whom sepsis could be predicted 5 to 6 hours in advance, and XGBoost successfully predicted 60% (24/40) of them at the 10% precision level. The calibration curve showed that the predictions from XGBoost consistently overestimated the risk, whereas the predictions after recalibration lay snugly around the diagonal ([Figure 3](#)). The DCA demonstrated that XGBoost had a positive net benefit in clinical use for threshold probability across the threshold probabilities of 0 to 0.6 ([Figure 3](#)).

**Table 2.** Summary of model performance on the test set for Extreme Gradient Boosting (XGBoost) and logistic regression.

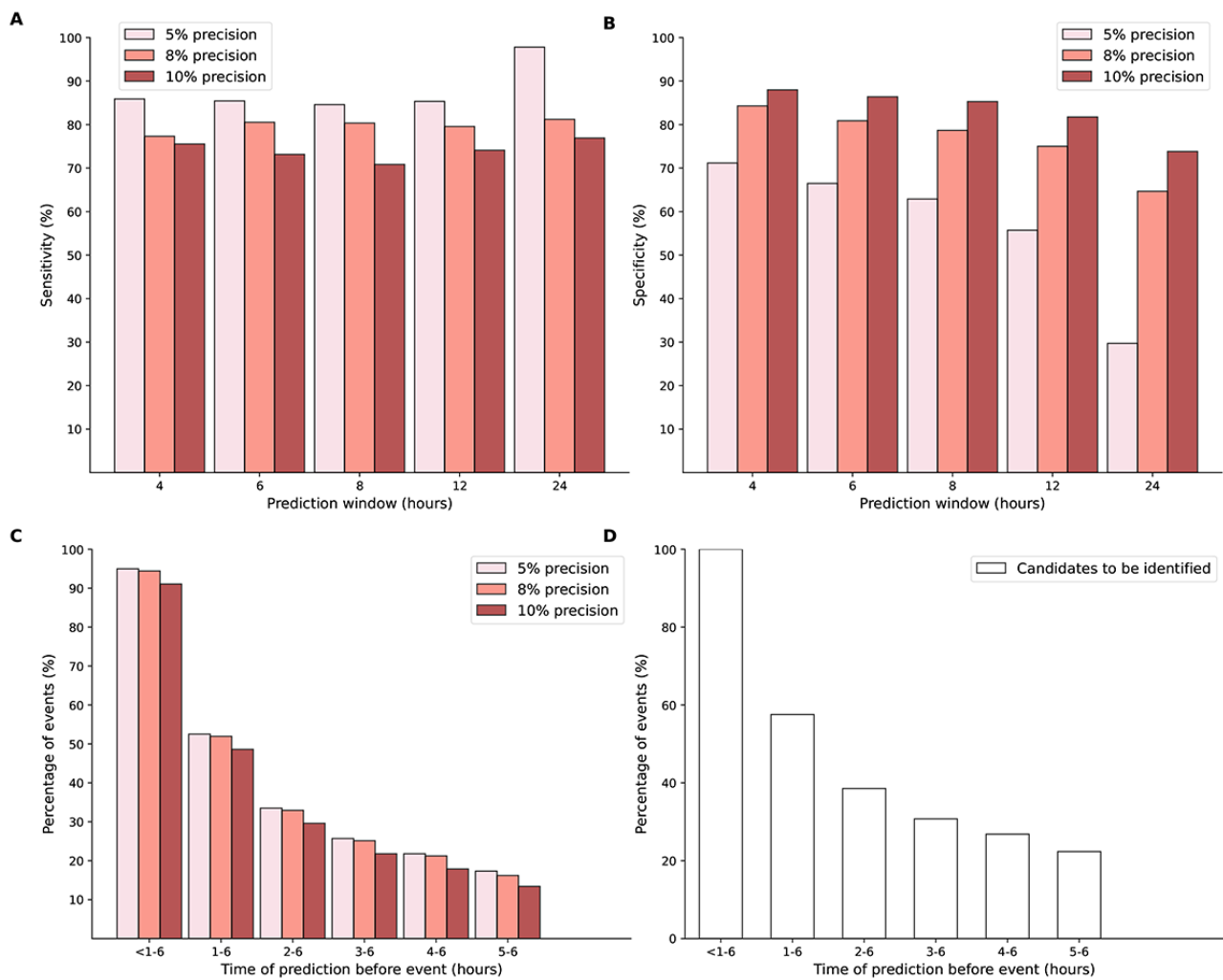
Performance metric	Value at 4 hours (95% CI)	Value at 6 hours (95% CI)	Value at 8 hours (95% CI)	Value at 12 hours (95% CI)	Value at 24 hours (95% CI)
<b>XGBoost</b>					
AUROC <sup>a</sup>	0.88 (0.86-0.90)	0.87 (0.85-0.89)	0.86 (0.84-0.87)	0.84 (0.83-0.86)	0.83 (0.81-0.84)
AUPRC <sup>b</sup>	0.27 (0.23-0.31)	0.27 (0.23-0.31)	0.26 (0.23-0.30)	0.25 (0.22-0.28)	0.23 (0.20-0.26)
ACE <sup>c</sup>	0.35 (0.32-0.37)	0.33 (0.31-0.35)	0.32 (0.30-0.35)	0.32 (0.30-0.34)	0.30 (0.28-0.32)
<b>Logistic regression</b>					
AUROC	0.84 (0.82-0.86)	0.83 (0.81-0.85)	0.81 (0.79-0.83)	0.79 (0.77-0.80)	0.76 (0.74-0.77)
AUPRC	0.18 (0.15-0.22)	0.18 (0.15-0.21)	0.18 (0.15-0.21)	0.17 (0.14-0.20)	0.16 (0.14-0.18)
ACE	0.45 (0.44-0.45)	0.44 (0.44-0.45)	0.44 (0.44-0.44)	0.44 (0.43-0.44)	0.42 (0.42-0.43)

<sup>a</sup>AUROC: area under the receiver operating characteristic curve.

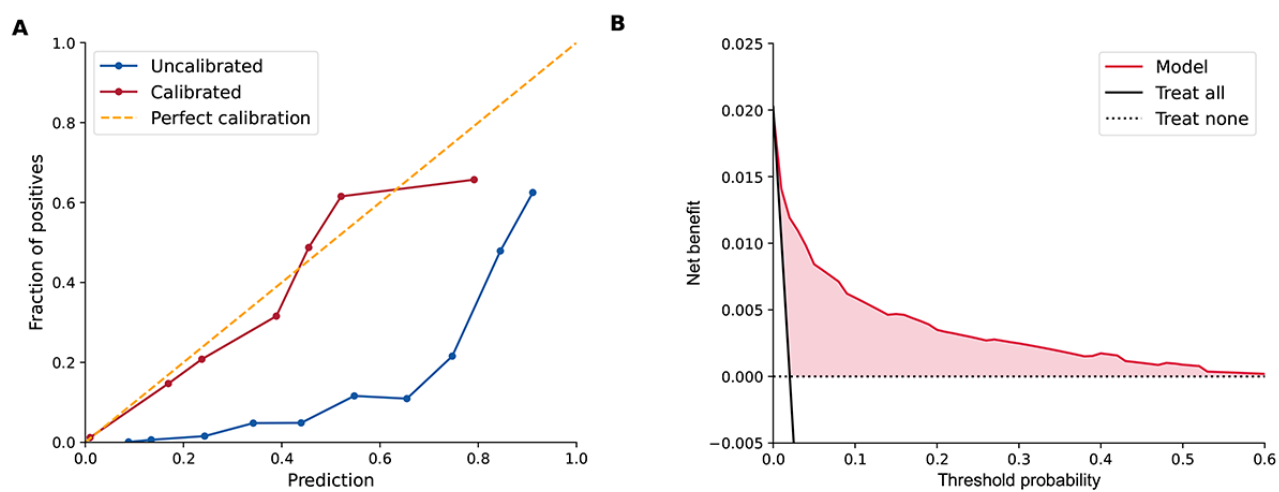
<sup>b</sup>AUPRC: area under the precision-recall curve.

<sup>c</sup>ACE: average calibration error.

**Figure 2.** Time-step-level and outcome-level sensitivity and specificity by pairs of precision level (5%, 8%, and 10%) and prediction window for the Extreme Gradient Boosting (XGBoost) model. (A) Time-step-level sensitivity. (B) Time-step-level specificity. (C) Outcome-level sensitivity. (D) The proportion of (candidate) adverse events to be identified within each window.



**Figure 3.** Calibration and clinical utility of the Extreme Gradient Boosting (XGBoost) model. (A) Calibration curves before and after calibration. (B) Decision curve.

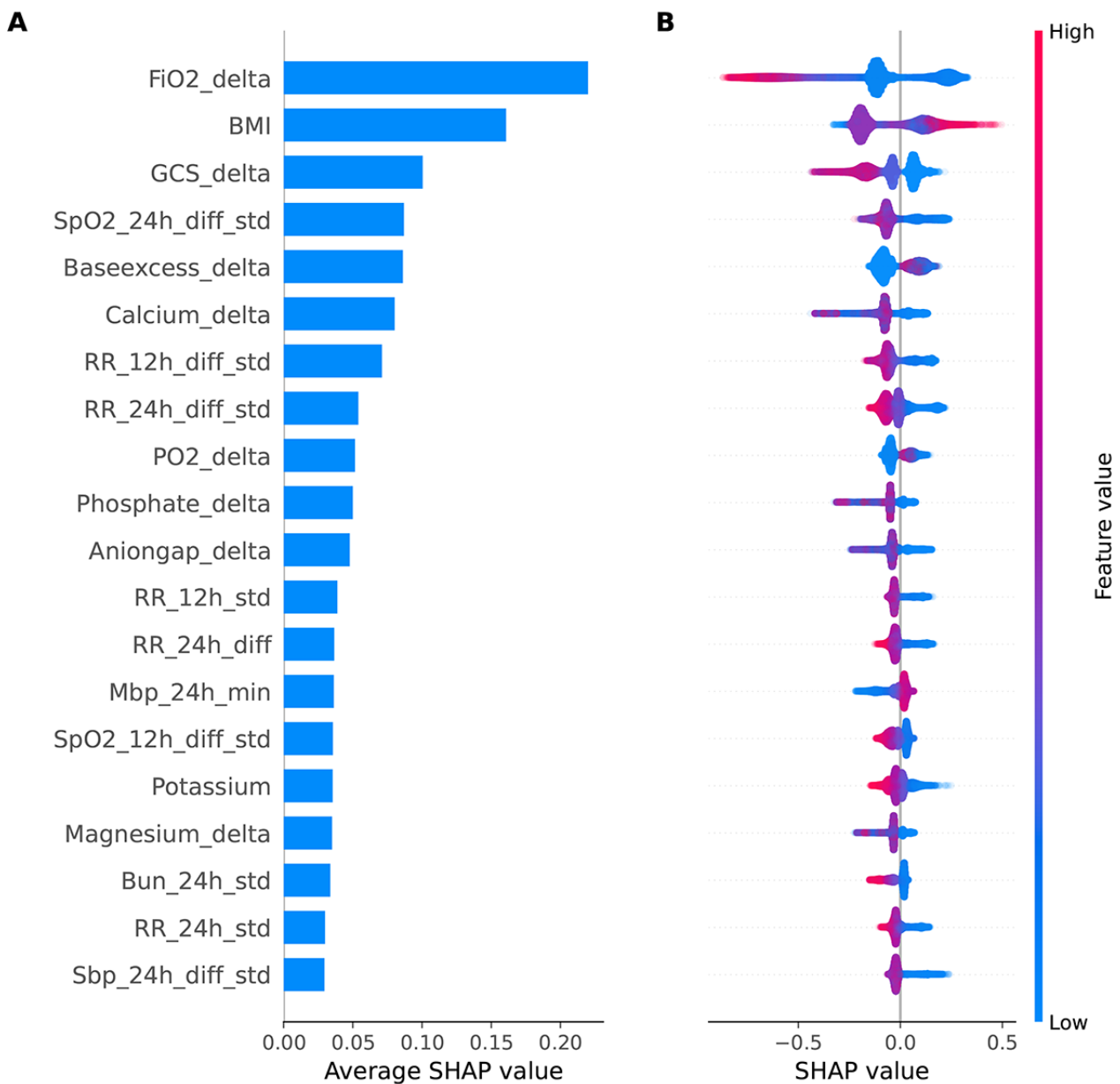


**SHAP Algorithm**

When considering the relative importance of each feature in the model, we found that the latest measurement time gap of fraction of inspired oxygen (FiO<sub>2</sub>) had the greatest impact on the

predictions, followed by BMI (Figure 4). Patients with a shorter measurement time gap of FiO<sub>2</sub> or a higher BMI had an increased risk of sepsis. For time series variables of SD, differential SD, and the difference between maximum and minimum values of a feature, low values increased the risk of sepsis.

**Figure 4.** Bar plots showing (A) overall impacts of the top 20 features and (B) beeswarm plots showing impacts of the top 20 features across all patients. BMI: body mass index; Bun: blood urea nitrogen; delta: the latest measurement time gap; diff\_std: differential standard deviation; diff: the difference between maximum and minimum values; FiO<sub>2</sub>: fraction of inspired oxygen; GCS: Glasgow Coma Scale; Mbp: mean blood pressure; PO<sub>2</sub>: arterial partial pressure of oxygen; RR: respiratory rate; Sbp: systolic blood pressure; SHAP: Shapley additive explanation; SpO<sub>2</sub>: saturation of peripheral oxygen; std: standard deviation.



## Discussion

### Principal Findings

In this study, we developed an XGBoost risk prediction model to predict sepsis onset among trauma patients admitted to the ICU with a temporal resolution of 1 hour. This model achieved an AUROC ranging from 0.83 to 0.88 at the 4-to-24-hour prediction window. It predicted 73% (386/529) of sepsis-positive time-steps and 91% (163/179) of sepsis events in the subsequent 6 hours with a ratio of 9 false alerts for every true alert. Furthermore, the model achieved better discriminative and calibration performance than a traditional logistic regression model. However, this finding remains to be validated in other data sets; the classical logistic regression might be suboptimal compared with the XGBoost model.

Wong et al [10] recently reported that the widely applied ESM only identified sepsis before onset in 33% of patients, whereas our model identified up to 91% (163/179) of patients who developed sepsis in the subsequent 6 hours at 10% precision. To our knowledge, the XGBoost model in our study has better discrimination performance (with an AUROC of 0.87) than most previously published models that have been developed for real-time prediction of sepsis in the general ICU setting. Nemati et al [35] achieved an AUROC of 0.85 with a modified Weibull-Cox proportional hazards model for predicting sepsis 6 hours in advance, and Yang et al [28] achieved similar performance, also with the XGBoost algorithm. Kim et al [36] recently developed a type of deep learning model to predict sepsis that had higher discrimination performance than our model, with an AUROC of 0.91. However, their model could

be seen as a complex black box due to its lack of interpretability, which might limit its acceptance among clinicians. Moreover, deep learning models like neural networks usually have a large number of parameters to estimate and have poor generalizability without sufficient training data, and it takes longer to train them than XGBoost [37]. The random forest model is another widely applied machine learning approach, but XGBoost might be a better option for imbalanced data sets, such as the one used in our study [38,39].

In addition to AUROC, a commonly reported measure of discriminative performance, we report AUPRC results for our model. This is more informative in class-imbalanced situations [22], such as sepsis prediction. Our model had an AUPRC of 0.27, which indicates low precision across a wide range of sensitivities in this extremely imbalanced data set. We found that the model achieved higher AUROC and AUPRC with shorter prediction windows. This could be attributed to the fact that a decreasing prediction window improved the timeliness of information, which boosted the predictive performance of the model. Moreover, we report calibration performance in addition to the commonly reported discriminative performance [40]. Calibration evaluates the agreement between the estimated and true risk of an outcome [41], which is important when a model is designed to make predictions at an individual level. Here, our model had an ACE of 0.33 before using isotonic regression calibration, which suggests that the model overestimated the risk of sepsis. However, model calibration decreased as the prediction window shortened, which might be associated with a decreasing number of positive steps due to the reduction of the prediction window. Several studies have reported a similar trend for AUROC across different prediction windows but have not reported changes in AUPRC or calibration [35,36]. Most importantly, a model with good discrimination and calibration performance does not necessarily have high clinical value [42]. Hence, DCA was used to assess the clinical utility of the model, and this showed a positive net benefit, suggesting that the model could help to inform timely treatment before sepsis onset in clinical practice. As the net benefit takes into account both true positives and false positives, the model with a net benefit is therefore worth choosing irrespective of the size or statistical significance of the benefit [42]. However, our model is not a practical tool at present, and we plan to develop a handy risk prediction tool by integrating the model into electronic health records for early identification of sepsis among trauma patients.

The matter of model applicability has not been well addressed in previous studies. In this study, we evaluated the time-step-level sensitivity and specificity of the model at different degrees of precision. The precision explicitly shows the number of false positives that the clinician encountered to

identify one true positive episode or case. However, the sequential nature of making predictions determines the total number of positive steps; this does not directly correspond to the total number of patients with sepsis. Multiple positive time steps may be associated with a single sepsis episode. In fact, one positive prediction in the prediction window was enough to attract the attention of a clinician to make further decisions. Therefore, we calculated the outcome-level sensitivity (ie, the percentage of all sepsis episodes that had at least one correct prediction within a fixed time window before sepsis onset) to show the ability of the model to identify the percentage of true positive patients [43,44]. Furthermore, some previous studies have screened patients based on length of stay in the hospital, which might influence the generalizability and implementation of the model in a prospective setting [22,45].

Through SHAP analysis, we found that obese trauma patients were at an increased risk of sepsis. Obesity is associated with altered cellular immunity, increased use of central venous catheters because of difficulties with gaining peripheral access, and inadequate antibiotic dosing, all of which increase the risk of sepsis [46,47]. Moreover, obesity is associated with comorbidities like diabetes and hypertension, which have been identified as risk factors for sepsis [46,48]. Low values for time-series variables, such as differential SD of SpO<sub>2</sub> and differential SD of respiratory rate, were associated with an increased risk of sepsis. One possible explanation was that most patients developed sepsis in a short time after admission. We compared the top 20 variables in the sorted SHAP value diagram in our model for critical trauma patients with those from other models developed for general critical patients and found that BMI ranked second for trauma patients but was not in the top 20 for general critical patients [28,49]. Contrarily, age ranked 14th for general critical patients but was not in the top 20 for trauma patients (Figure 4) [49].

### Limitations

This study has several limitations. First, though our model has shown good performance and clinical utility, it needs to be further validated at other medical centers. Second, the Injury Severity Score was not used for model development, even though this score is commonly used for assessing injury severity and might contain predictive information for sepsis in the trauma population. However, the Injury Severity Score is not available in the MIMIC database, and it is not an objective metric [50].

### Conclusions

In summary, an XGBoost model achieved high performance in both discrimination and calibration for continuous prediction of sepsis onset in the next 6 hours among trauma patients. Furthermore, the model was clinically useful and had a positive net benefit across the threshold probability.

### Acknowledgments

XW was supported by the National Science and Technology Major Project of China (grant 2018ZX10713001-007) and by the Three-year Public Health System Construction Program of Shanghai, China (grant GWV-10.2-YQ36). WY was supported by the 13th Five-Year Plan Foundation of Jiangsu Province for Medical Key Talents (grant ZDRCA2016099) and by the National Major Scientific Research Instruments and Equipment Development Project of the National Natural Science Foundation of China (grant 81927808).



## Data Availability

The Medical Information Mart for Intensive Care IV (MIMIC IV) data set used for analyses in this study has been published and made publicly available [18]. The code used in this study is largely based on the *numpy*, *pandas*, and *scikit-learn* libraries in python and is available from the corresponding authors upon reasonable request.

## Authors' Contributions

XW conceived, designed, and supervised the study. FX and CS participated in data wrangling. JL carried out the data analysis. JL and FX prepared the first draft of the manuscript. WY and XW interpreted the findings and commented on and revised drafts of the manuscript. All authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

List of raw variables (n=42) extracted from the MIMIC IV database. MIMIC IV: Medical Information Mart for Intensive Care IV.

[\[DOCX File , 19 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

The definition of the onset time of sepsis.

[\[DOCX File , 14 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

Details of the feature engineering.

[\[DOCX File , 18 KB-Multimedia Appendix 3\]](#)

## Multimedia Appendix 4

Receiver operating characteristic (ROC) curves (A) and precision versus the sensitivity (PR) curves (B) for extreme gradient boosting (XGBoost) and logistic regression. AUROC: area under the receiver operating characteristic curve; AUPRC: area under the precision-recall curve.

[\[PDF File \(Adobe PDF File\), 37 KB-Multimedia Appendix 4\]](#)

## References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016 Feb 23;315(8):801-810 [[FREE Full text](#)] [doi: [10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287)] [Medline: [26903338](https://pubmed.ncbi.nlm.nih.gov/26903338/)]
2. Eguia E, Bunn C, Kulshrestha S, Markossian T, Durazo-Arvizu R, Baker MS, et al. Trends, cost, and mortality from sepsis after trauma in the United States: An evaluation of the national inpatient sample of hospitalizations, 2012-2016. *Crit Care Med* 2020 Sep;48(9):1296-1303 [[FREE Full text](#)] [doi: [10.1097/CCM.0000000000004451](https://doi.org/10.1097/CCM.0000000000004451)] [Medline: [32590387](https://pubmed.ncbi.nlm.nih.gov/32590387/)]
3. Lindner HA, Balaban, Sturm T, Weiß C, Thiel M, Schneider-Lindner V. An algorithm for systemic inflammatory response syndrome criteria-based prediction of sepsis in a polytrauma cohort. *Crit Care Med* 2016 Dec;44(12):2199-2207. [doi: [10.1097/CCM.0000000000001955](https://doi.org/10.1097/CCM.0000000000001955)] [Medline: [27441906](https://pubmed.ncbi.nlm.nih.gov/27441906/)]
4. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med* 2017 Mar;43(3):304-377 [[FREE Full text](#)] [doi: [10.1007/s00134-017-4683-6](https://doi.org/10.1007/s00134-017-4683-6)] [Medline: [28101605](https://pubmed.ncbi.nlm.nih.gov/28101605/)]
5. Ferrer R, Martin-Loeches I, Phillips G, Osborn TM, Townsend S, Dellinger RP, et al. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Crit Care Med* 2014 Aug;42(8):1749-1755. [doi: [10.1097/CCM.0000000000000330](https://doi.org/10.1097/CCM.0000000000000330)] [Medline: [24717459](https://pubmed.ncbi.nlm.nih.gov/24717459/)]
6. Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med* 2006 Jun;34(6):1589-1596. [doi: [10.1097/01.CCM.0000217961.75225.E9](https://doi.org/10.1097/01.CCM.0000217961.75225.E9)] [Medline: [16625125](https://pubmed.ncbi.nlm.nih.gov/16625125/)]
7. Vincent J. The clinical challenge of sepsis identification and monitoring. *PLoS Med* 2016 May;13(5):e1002022 [[FREE Full text](#)] [doi: [10.1371/journal.pmed.1002022](https://doi.org/10.1371/journal.pmed.1002022)] [Medline: [27187803](https://pubmed.ncbi.nlm.nih.gov/27187803/)]

8. de Grooth H, Postema J, Loer SA, Parienti J, Oudemans-van Straaten HM, Girbes AR. Unexplained mortality differences between septic shock trials: a systematic analysis of population characteristics and control-group mortality rates. *Intensive Care Med* 2018 Mar;44(3):311-322 [[FREE Full text](#)] [doi: [10.1007/s00134-018-5134-8](https://doi.org/10.1007/s00134-018-5134-8)] [Medline: [29546535](#)]
9. Goh KH, Wang L, Yeow AYK, Poh H, Li K, Yeow JYL, et al. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nat Commun* 2021 Jan 29;12(1):711 [[FREE Full text](#)] [doi: [10.1038/s41467-021-20910-4](https://doi.org/10.1038/s41467-021-20910-4)] [Medline: [33514699](#)]
10. Wong A, Otlés E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med* 2021 Aug 01;181(8):1065-1070 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2021.2626](https://doi.org/10.1001/jamainternmed.2021.2626)] [Medline: [34152373](#)]
11. Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, et al. Machine learning-based early warning systems for clinical deterioration: Systematic scoping review. *J Med Internet Res* 2021 Feb 04;23(2):e25187 [[FREE Full text](#)] [doi: [10.2196/25187](https://doi.org/10.2196/25187)] [Medline: [33538696](#)]
12. Joshi M, Ashrafian H, Arora S, Khan S, Cooke G, Darzi A. Digital alerting and outcomes in patients with sepsis: Systematic review and meta-analysis. *J Med Internet Res* 2019 Dec 20;21(12):e15166 [[FREE Full text](#)] [doi: [10.2196/15166](https://doi.org/10.2196/15166)] [Medline: [31859672](#)]
13. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med* 2020 Mar;46(3):383-400 [[FREE Full text](#)] [doi: [10.1007/s00134-019-05872-y](https://doi.org/10.1007/s00134-019-05872-y)] [Medline: [31965266](#)]
14. Moor M, Rieck B, Horn M, Jutzeler CR, Borgwardt K. Early prediction of sepsis in the ICU using machine learning: A systematic review. *Front Med (Lausanne)* 2021;8:607952 [[FREE Full text](#)] [doi: [10.3389/fmed.2021.607952](https://doi.org/10.3389/fmed.2021.607952)] [Medline: [34124082](#)]
15. Chalya PL, Gilyoma JM, Dass RM, Mchembe MD, Matasha M, Mabula JB, et al. Trauma admissions to the intensive care unit at a reference hospital in Northwestern Tanzania. *Scand J Trauma Resusc Emerg Med* 2011 Oct 24;19(1):61 [[FREE Full text](#)] [doi: [10.1186/1757-7241-19-61](https://doi.org/10.1186/1757-7241-19-61)] [Medline: [22024353](#)]
16. D'Abbondanza JA, Shahrokhi S. Burn infection and burn sepsis. *Surg Infect (Larchmt)* 2021 Feb;22(1):58-64. [doi: [10.1089/sur.2020.102](https://doi.org/10.1089/sur.2020.102)] [Medline: [32364824](#)]
17. Mankowski RT, Anton SD, Ghita GL, Brumback B, Darden DB, Bihorac A, et al. Older adults demonstrate biomarker evidence of the persistent inflammation, immunosuppression, and catabolism syndrome (PICS) after sepsis. *J Gerontol A Biol Sci Med Sci* 2022 Jan 07;77(1):188-196 [[FREE Full text](#)] [doi: [10.1093/gerona/glab080](https://doi.org/10.1093/gerona/glab080)] [Medline: [33721883](#)]
18. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV (version 1.0). *PhysioNet*. 2021. URL: <https://physionet.org/content/mimiciv/1.0/> [accessed 2023-03-16]
19. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data* 2023 Jan 03;10(1):1 [[FREE Full text](#)] [doi: [10.1038/s41597-022-01899-x](https://doi.org/10.1038/s41597-022-01899-x)] [Medline: [36596836](#)]
20. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of clinical criteria for sepsis: For the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016 Feb 23;315(8):762-774 [[FREE Full text](#)] [doi: [10.1001/jama.2016.0288](https://doi.org/10.1001/jama.2016.0288)] [Medline: [26903335](#)]
21. Moor M, Horn M, Rieck B, Roqueiro D, Borgwardt K. Temporal convolutional networks and dynamic time warping can drastically improve the early prediction of sepsis. *ArXiv*. Preprint posted online Feb 7, 2019. [[FREE Full text](#)]
22. Lauritsen SM, Thiesson B, Jørgensen MJ, Riis AH, Espelund US, Weile JB, et al. The framing of machine learning risk prediction models illustrated by evaluation of sepsis in general wards. *NPJ Digit Med* 2021 Nov 15;4(1):158 [[FREE Full text](#)] [doi: [10.1038/s41746-021-00529-x](https://doi.org/10.1038/s41746-021-00529-x)] [Medline: [34782696](#)]
23. Thorsen-Meyer H, Nielsen AB, Nielsen AP, Kaas-Hansen BS, Toft P, Schierbeck J, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Digit Health* 2020 Apr;2(4):e179-e191 [[FREE Full text](#)] [doi: [10.1016/S2589-7500\(20\)30018-2](https://doi.org/10.1016/S2589-7500(20)30018-2)] [Medline: [33328078](#)]
24. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
25. Kausch SL, Moorman JR, Lake DE, Keim-Malpass J. Physiological machine learning models for prediction of sepsis in hospitalized adults: An integrative review. *Intensive Crit Care Nurs* 2021 Aug;65:103035. [doi: [10.1016/j.iccn.2021.103035](https://doi.org/10.1016/j.iccn.2021.103035)] [Medline: [33875337](#)]
26. Deng H, Sun M, Wang Y, Zeng J, Yuan T, Li T, et al. Evaluating machine learning models for sepsis prediction: A systematic review of methodologies. *iScience* 2022 Jan 21;25(1):103651 [[FREE Full text](#)] [doi: [10.1016/j.isci.2021.103651](https://doi.org/10.1016/j.isci.2021.103651)] [Medline: [35028534](#)]
27. Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N. Taking the human out of the loop: a review of bayesian optimization. *Proc IEEE* 2016 Jan;104(1):148-175 [[FREE Full text](#)] [doi: [10.1109/jproc.2015.2494218](https://doi.org/10.1109/jproc.2015.2494218)]

28. Yang M, Liu C, Wang X, Li Y, Gao H, Liu X, et al. An explainable artificial intelligence predictor for early detection of sepsis. *Crit Care Med* 2020 Nov;48(11):e1091-e1096. [doi: [10.1097/CCM.0000000000004550](https://doi.org/10.1097/CCM.0000000000004550)] [Medline: [32885937](https://pubmed.ncbi.nlm.nih.gov/32885937/)]
29. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. 2005 Presented at: ICML-2005, the 22nd International Conference on Machine Learning; August 7-11, 2005; Bonn, Germany. [doi: [10.1145/1102351.1102430](https://doi.org/10.1145/1102351.1102430)]
30. Zadrozny B, Elkan C. Transforming classifier scores into accurate multiclass probability estimates. In: *KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2002 Presented at: KDD02: The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; July 23-26, 2002; Edmonton, AB. [doi: [10.1145/775047.775151](https://doi.org/10.1145/775047.775151)]
31. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ* 2020 May 20;369:m1501 [FREE Full text] [doi: [10.1136/bmj.m1501](https://doi.org/10.1136/bmj.m1501)] [Medline: [32434791](https://pubmed.ncbi.nlm.nih.gov/32434791/)]
32. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and interpreting decision curve analysis: A guide for investigators. *Eur Urol* 2018 Dec;74(6):796-804 [FREE Full text] [doi: [10.1016/j.eururo.2018.08.038](https://doi.org/10.1016/j.eururo.2018.08.038)] [Medline: [30241973](https://pubmed.ncbi.nlm.nih.gov/30241973/)]
33. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems 30*. 2017 Presented at: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017); December 4-9, 2017; Long Beach, CA p. 4765-4774 URL: <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
34. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020 Jan;2(1):56-67 [FREE Full text] [doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9)] [Medline: [32607472](https://pubmed.ncbi.nlm.nih.gov/32607472/)]
35. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Crit Care Med* 2018 Apr;46(4):547-553 [FREE Full text] [doi: [10.1097/CCM.0000000000002936](https://doi.org/10.1097/CCM.0000000000002936)] [Medline: [29286945](https://pubmed.ncbi.nlm.nih.gov/29286945/)]
36. Kim JK, Ahn W, Park S, Lee S, Kim L. Early prediction of sepsis onset using neural architecture search based on genetic algorithms. *Int J Environ Res Public Health* 2022 Feb 18;19(4):2349 [FREE Full text] [doi: [10.3390/ijerph19042349](https://doi.org/10.3390/ijerph19042349)] [Medline: [35206537](https://pubmed.ncbi.nlm.nih.gov/35206537/)]
37. Han X, Zhang Z, Ding N, Gu Y, Liu X, Huo Y, et al. Pre-trained models: Past, present and future. *AI Open* 2021;2:225-250 [FREE Full text] [doi: [10.1016/j.aiopen.2021.08.002](https://doi.org/10.1016/j.aiopen.2021.08.002)]
38. Didavi A, Agbokpanzo R, Agbomahena M. Comparative study of decision tree, random forest and XGBoost performance in forecasting the power output of a photovoltaic. 2021 Presented at: 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART); December 8-10, 2021; Paris/Créteil, France. [doi: [10.1109/biosmart54244.2021.9677566](https://doi.org/10.1109/biosmart54244.2021.9677566)]
39. Mishra S. Handling imbalanced data: SMOTE vs. random undersampling. *Int Res J Eng Technol* 2017;4(8):317-320 [FREE Full text]
40. Giacobbe DR, Signori A, Del Puente F, Mora S, Carmisciano L, Briano F, et al. Early detection of sepsis with machine learning techniques: A brief clinical perspective. *Front Med (Lausanne)* 2021;8:617486 [FREE Full text] [doi: [10.3389/fmed.2021.617486](https://doi.org/10.3389/fmed.2021.617486)] [Medline: [33644097](https://pubmed.ncbi.nlm.nih.gov/33644097/)]
41. Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. *J Am Med Inform Assoc* 2020 Apr 01;27(4):621-633 [FREE Full text] [doi: [10.1093/jamia/ocz228](https://doi.org/10.1093/jamia/ocz228)] [Medline: [32106284](https://pubmed.ncbi.nlm.nih.gov/32106284/)]
42. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019;3:18 [FREE Full text] [doi: [10.1186/s41512-019-0064-7](https://doi.org/10.1186/s41512-019-0064-7)] [Medline: [31592444](https://pubmed.ncbi.nlm.nih.gov/31592444/)]
43. Pimentel MAF, Redfern OC, Malycha J, Meredith P, Prytherch D, Briggs J, et al. Detecting deteriorating patients in the hospital: Development and validation of a novel scoring system. *Am J Respir Crit Care Med* 2021 Jul 01;204(1):44-52 [FREE Full text] [doi: [10.1164/rccm.202007-2700OC](https://doi.org/10.1164/rccm.202007-2700OC)] [Medline: [33525997](https://pubmed.ncbi.nlm.nih.gov/33525997/)]
44. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019 Aug;572(7767):116-119 [FREE Full text] [doi: [10.1038/s41586-019-1390-1](https://doi.org/10.1038/s41586-019-1390-1)] [Medline: [31367026](https://pubmed.ncbi.nlm.nih.gov/31367026/)]
45. Reyna MA, Josef CS, Jeter R, Shashikumar SP, Westover MB, Nemati S, et al. Early prediction of sepsis from clinical data: The PhysioNet/Computing in Cardiology Challenge 2019. *Crit Care Med* 2020 Feb;48(2):210-217 [FREE Full text] [doi: [10.1097/CCM.0000000000004145](https://doi.org/10.1097/CCM.0000000000004145)] [Medline: [31939789](https://pubmed.ncbi.nlm.nih.gov/31939789/)]
46. Anderson MR, Shashaty MGS. Impact of obesity in critical illness. *Chest* 2021 Dec;160(6):2135-2145 [FREE Full text] [doi: [10.1016/j.chest.2021.08.001](https://doi.org/10.1016/j.chest.2021.08.001)] [Medline: [34364868](https://pubmed.ncbi.nlm.nih.gov/34364868/)]
47. Bochicchio GV, Joshi M, Bochicchio K, Nehman S, Tracy JK, Scalea TM. Impact of obesity in the critically ill trauma patient: a prospective study. *J Am Coll Surg* 2006 Oct;203(4):533-538. [doi: [10.1016/j.jamcollsurg.2006.07.001](https://doi.org/10.1016/j.jamcollsurg.2006.07.001)] [Medline: [17000398](https://pubmed.ncbi.nlm.nih.gov/17000398/)]
48. Locham S, Naazie I, Canner J, Siracuse J, Al-Nouri O, Malas M. Incidence and risk factors of sepsis in hemodialysis patients in the United States. *J Vasc Surg* 2021 Mar;73(3):1016-1021.e3 [FREE Full text] [doi: [10.1016/j.jvs.2020.06.126](https://doi.org/10.1016/j.jvs.2020.06.126)] [Medline: [32707386](https://pubmed.ncbi.nlm.nih.gov/32707386/)]

49. Zhao X, Shen W, Wang G. Early prediction of sepsis based on machine learning algorithm. *Comput Intell Neurosci* 2021;2021:6522633 [FREE Full text] [doi: [10.1155/2021/6522633](https://doi.org/10.1155/2021/6522633)] [Medline: [34675971](https://pubmed.ncbi.nlm.nih.gov/34675971/)]
50. Rutledge R. The Injury Severity Score is unable to differentiate between poor care and severe injury. *J Trauma* 1996 Jun;40(6):944-950. [doi: [10.1097/00005373-199606000-00013](https://doi.org/10.1097/00005373-199606000-00013)] [Medline: [8656481](https://pubmed.ncbi.nlm.nih.gov/8656481/)]

## Abbreviations

**ACE:** average calibration error  
**AUPRC:** area under the precision-recall curve  
**AUROC:** area under the receiver operating characteristic curve  
**DCA:** decision curve analysis  
**ESM:** Epic Sepsis Model  
**FiO<sub>2</sub>:** fraction of inspired oxygen  
**ICU:** intensive care unit  
**MIMIC IV:** Medical Information Mart for Intensive Care IV  
**SHAP:** Shapley additive explanation  
**SOFA:** Sequential Organ Failure Assessment  
**SpO<sub>2</sub>:** saturation of peripheral oxygen  
**XGBoost:** Extreme Gradient Boosting

*Edited by A Mavragani; submitted 05.09.22; peer-reviewed by L Li, E Steyerberg; comments to author 08.10.22; revised version received 21.10.22; accepted 23.02.23; published 31.03.23*

*Please cite as:*

*Li J, Xi F, Yu W, Sun C, Wang X*

*Real-Time Prediction of Sepsis in Critical Trauma Patients: Machine Learning–Based Modeling Study*

*JMIR Form Res* 2023;7:e42452

URL: <https://formative.jmir.org/2023/1/e42452>

doi: [10.2196/42452](https://doi.org/10.2196/42452)

PMID: [37000488](https://pubmed.ncbi.nlm.nih.gov/37000488/)

©Jiang Li, Fengchan Xi, Wenkui Yu, Chuanrui Sun, Xiling Wang. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.03.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.