

Machine learning for prediction of in-ICU mortality in sepsis: an observational study using the MIMIC IV database

Sitong Liu¹, Lingmin Gong¹, Weiwei Hu¹, Jiaxin Cai¹, Yuhui Yang¹, Shiyu Chen¹, Baibing Mi¹, Yaling Zhao¹, Leilei Pei¹, Fangyao Chen^{1,2*}

¹Department of Epidemiology and Biostatistics, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China

²Department of Radiology, First Affiliate Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, China

Submitted: 10 July 2024; **Accepted:** 12 December 2024

Online publication: 20 April 2025

Arch Med Sci

DOI: <https://doi.org/10.5114/aoms/197307>

Copyright © 2025 Termedia & Banach

*Corresponding author:

Fangyao Chen
Department of
Epidemiology and
Biostatistics
School of Public Health
Xi'an Jiaotong
University Health
Science Center
Department of Radiology
First Affiliate Hospital
of Xi'an Jiaotong University
Xi'an, Shaanxi, China
E-mail: chenfy@xjtu.edu.cn

Abstract

Introduction: The aim of the study was to identify the key risk factors influencing in-intensive care unit (ICU) mortality of patients with sepsis and develop prognosis prediction models for culture-positive sepsis (CPS) and culture-negative sepsis (CNS) patients.

Material and methods: Data were extracted from the MIMIC-IV database, which included 9288 patients with sepsis. The whole sample was divided into CPS (6622 patients) and CNS groups (2666 patients). We established six machine learning models – DT, RF, NB, XGB, GBDT, and NNET – to predict in-ICU death for all study samples, as well as for CPS and CNS subgroups. Model performance was assessed using AUC, accuracy, sensitivity, and specificity. SHapley Additive exPlanations (SHAP) values were used to explain the effect of variables on model results.

Results: The in-ICU mortality rate was 54.58% for the whole study sample; the difference in in-ICU mortality between the CPS (55.19%) and CNS (53.04%) groups was not statistically significant. The main significant influential factors identified included Charlson Comorbidity Index (CCI), number of days in hospital, Glasgow Coma Scale (GCS), older age, and total bilirubin (TBil). The XGB model performed best in the overall sample (AUC = 0.782), while the GBDT model was most effective for the CPS group (AUC = 0.7813) and the CNS group (AUC = 0.7582).

Conclusions: This study identified key risk factors for in-ICU death in patients with sepsis and highlighted differences in clinical characteristics between patients with CPS and CNS. These findings may contribute to the development of personalized treatment strategies and risk assessment, thereby improving the prognosis of septic patients, especially patients with CNS.

Key words: sepsis, in-intensive care unit mortality, machine learning, blood culture, interpretable learning.

Introduction

Sepsis is a life-threatening condition characterized by organ dysfunction resulting from a dysregulated host response to infection [1]. It remains a major public health challenge and is a leading cause of mortality

among critically ill patients in the intensive care unit (ICU) [2]. Sepsis affects more than 19 million people worldwide each year, with a persistently high mortality rate [3]. Effective prognosis improvement can be achieved by accurately identifying the pathogen and administering targeted antibiotics [4]. Rapid microbiological techniques, such as blood culture, play a crucial role in identifying sepsis pathogens [5]. These results can guide timely adjustments to antimicrobial therapies, enhancing patient outcomes, reducing hospital stays, and minimizing healthcare costs [5].

Most sepsis cases are found to have bacterial origins, according to existing studies [6]. Blood cultures are used to distinguish between culture-positive sepsis (CPS) and culture-negative sepsis (CNS), based on pathogen detection within 24 h before or after ICU admission [7]. CNS refers to sepsis cases where no causative organism is detected within this time frame, while CPS indicates the presence of identifiable pathogens [1]. The relationship between blood culture results and patient prognosis is debated, with mixed conclusions in current literature [2]. Kim *et al.* suggest a possible association between culture-negative sepsis and organ dysfunction, although the data do not indicate a significant association between blood culture results and sepsis mortality [8]. Conversely, Hazwani *et al.* reported lower mortality rates and reduced organ dysfunction among culture-negative sepsis patients [9]. Meanwhile, Baker *et al.* found that patients in the CPS and CNS groups exhibited similar symptoms upon admission and received comparable treatments [10].

Among studies to date, sepsis research has predominantly focused on CPS, with fewer investigations into CNS. However, CNS comprises 28–49% of sepsis cases [8], making it a significant subset. Due to the absence of clear pathogenic diagnoses, CNS patients often miss out on targeted antibiotic treatments, potentially resulting in higher morbidity, mortality, and prolonged hospital stays [11, 12]. Additionally, the empirical use of broad-spectrum antibiotics for CNS patients can elevate the risk of antibiotic resistance and complicate treatment [11]. CNS cases may involve diverse pathogens such as viruses, fungi, and parasites or even non-infectious factors such as autoimmune disorders and drug reactions [1]. Understanding the prognosis and underlying factors of CNS can shed light on sepsis etiology and mechanisms, paving the way for broader therapeutic strategies and improved patient management.

Several factors influence ICU mortality in sepsis patients, with blood culture results being just one aspect [1]. Accurate prediction of high-risk patients can aid clinicians in assessing disease severity, optimizing treatment, and reducing poor out-

comes [3]. In recent years, machine learning (ML) has been increasingly applied to sepsis research [13]. Persson *et al.* developed a high-performing ML algorithm for sepsis prediction using routine ICU data [14], while Yao *et al.* demonstrated that the XGB model excelled at predicting in-hospital mortality for postoperative sepsis patients [15]. Zhou *et al.* created an ML-based mortality prediction model for S-AKI patients, with the CatBoost model showing the best predictive results [16]. Yang *et al.* highlighted the superior predictive performance of XGB and random forest models for predicting sepsis onset [17].

Identifying high-risk subgroups is essential for managing CPS and CNS patients and improving their prognosis, with machine learning playing a key role in this process [14–17]. However, existing models have limitations, as most research focuses on CPS, with limited analysis of ICU mortality risk factors in CNS patients. This gap exists for several reasons. First, identifying the infection source is more straightforward in CPS, while it is often challenging in CNS cases [11, 12]. Second, positive bacterial cultures provide reliable results, whereas negative cultures can be affected by factors such as improper sampling or prior antibiotic use [1]. Consequently, CNS studies are more complex and may yield less consistent conclusions compared to CPS. Additionally, CPS treatment is more targeted, whereas CNS patients typically receive broad-spectrum antibiotics based on clinical judgment [1]. Lastly, CNS patients are more heterogeneous due to the involvement of various pathogens or non-infectious causes, complicating their analysis and treatment.

Given these limitations in mortality prediction for sepsis patients and the potential of machine learning in healthcare, there is a pressing need to develop predictive models that evaluate risk factors for both CPS and CNS patients. This study aimed to identify in-ICU mortality risk factors, compare the prognostic differences between CPS and CNS patients, and create predictive models using machine learning techniques.

Material and methods

Sources of data

This study was conducted based on clinical records of sepsis patients who meet the diagnostic criteria of the Sepsis-3 definition obtained from the Medical Information Mart for Intensive Care database-IV version 0.4 (MIMIC-IV v0.4). The MIMIC-IV database is a freely accessible critical care database released by the Laboratory for Computational Physiology at the Massachusetts Institute of Technology [18]. Detailed descriptions of this database are available in the literature [18].

Study population

In this study, sepsis was defined according to the Third International Consensus Definitions [1]. The selection of study samples was based on the following inclusion and exclusion criteria.

Inclusion criteria: (1) patients diagnosed with sepsis: diagnosed according to the Sepsis 3.0 criteria, i.e., infection + Sequential Organ Failure Assessment (SOFA) score ≥ 2 , or patients diagnosed with sepsis in discharge diagnosis according to ICD-9.0; (2) age ≥ 18 years old; (3) stayed in ICU for more than 24 h; (4) microbial cultures performed within 24 h before and after ICU admission.

Exclusion criteria: (1) patients diagnosed with sepsis after 24 h of ICU admission; (2) patients with multiple ICU admissions were only included based on their first ICU admission information; (3) patients with microbial culture results indicating fungal or viral infections were excluded.

A total of 9350 patients from MIMIC-IV met the inclusion and exclusion criteria. Among them, 62 patients lacked information on ICU ward type and were excluded. Consequently, a total of 9288 patients were included in the study, with 6622 (71.30%) cases in the CPS group and 2666 (28.70%) cases in the CNS group (Figure 1).

Study outcome

The primary outcome event of this study, based on its significance, is in-ICU death of the selected septic patients.

Data pre-processing

Variables with $\geq 30\%$ missingness were excluded from the analysis. Multiple imputation was performed on the remaining missing variables. To further validate the robustness of the interpola-

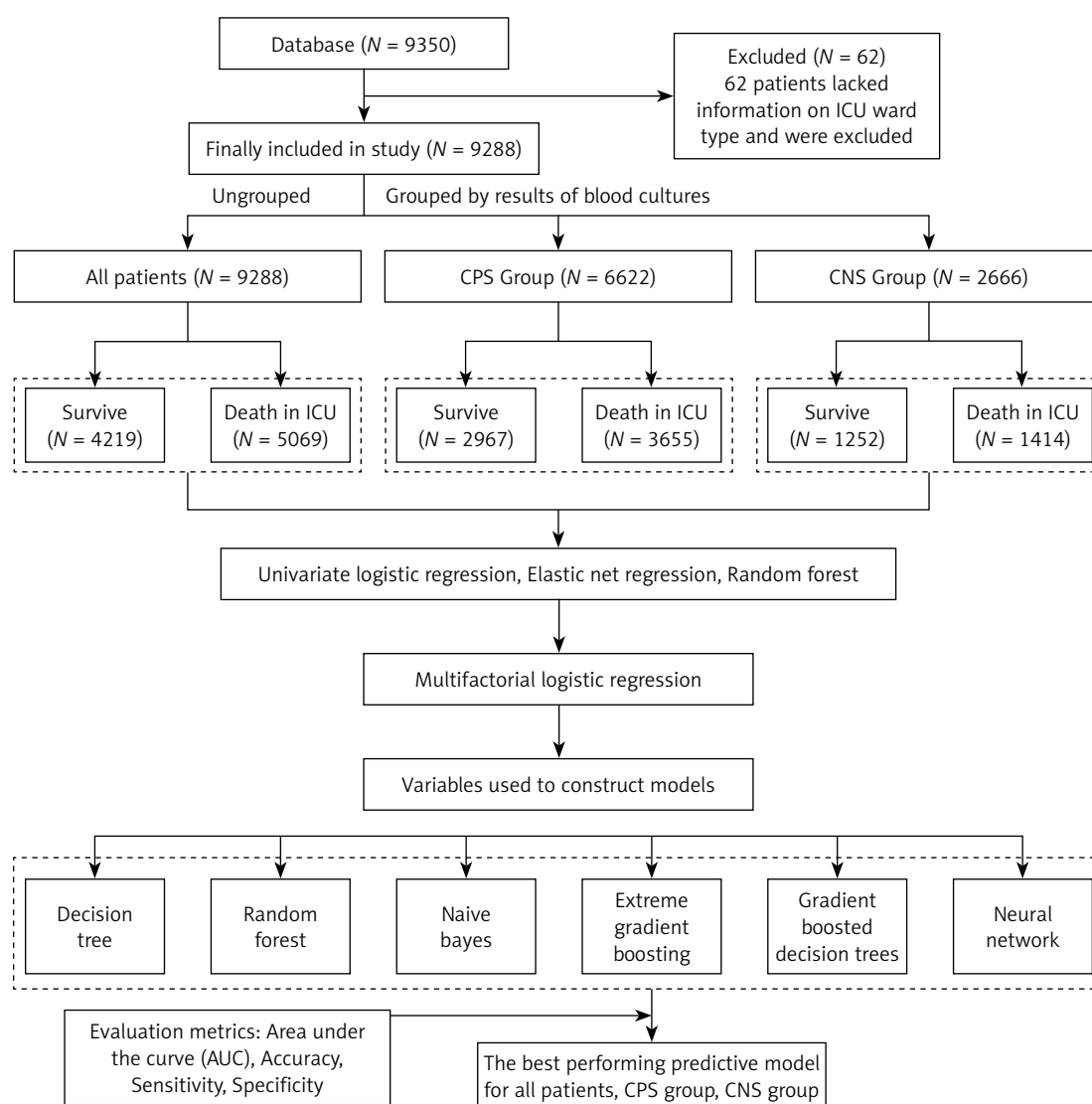


Figure 1. Study design

tion method, we used the “mice” package nested in the R programming language to generate five interpolation datasets using the multiple interpolation technique for variables containing missing values, and trained the machine learning model on each of the interpolation datasets separately to obtain the prediction results. Then the interpolated datasets were integrated using the pool function to achieve the final dataset for analysis. Multiple interpolation effectively reduces the error due to missing data by generating multiple datasets and integrating the interpolation uncertainty. We also transformed the continuous factor age in the database into an ordered factor age group. We recorded all cases of patients who had received vasopressin or ventilation or dialysis as 1. (The filtered variables and variable missingness rates are shown in Supplementary Table SI.)

Variable selection

To capture a broader range of variables associated with in-ICU death in the analysis, the following steps were taken. Firstly, we used three methods for factor screening in the training set [19], combining their results: (a) Univariate logistic regression (LR): each factor was analyzed individually, and factors with a two-tailed $p < 0.05$ were selected. (b) Elastic net regression: this regularized method optimized model complexity to identify important, sparse sets of factors [20]. (c) Random forest: feature importance was assessed using the Gini index to measure the contribution of each factor. Secondly, we fit a multifactorial logistic regression model with the selected variables. Then, the variables with $p < 0.05$ in multifactorial logistic regression were screened in conjunction with clinical significance. In addition, the above steps of screening variables were repeatedly conducted for the whole study sample, as well as the CPS and CNS subgroups. Finally, the valid variables identified from the overall sample, CPS group, and CNS group were input into the corresponding six machine learning models for each group.

Statistical analysis

Descriptive analysis was performed on all included patients. Continuous variables conforming to a normal distribution were described using means \pm standard deviations (SD). Skewed continuous variables were described using medians (interquartile ranges), and categorical variables were described using frequencies (proportions).

The datasets for the overall sample, CPS, and CNS were randomly sampled using the R programming language, and all were randomly divided into a training set and a test set in a ratio of 7 : 3. The training sets were used for selecting variables

and building models, and the test sets were used for validation.

Six machine learning methods – Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), Extreme gradient boosting (XGB), Gradient Boosted Decision Trees (GBDT), and Neural Network (NNET) models – were employed to establish predict models. The training and testing process involves 5-fold cross-validation. Subsequently, the accuracy, sensitivity, specificity, and area under the curve (AUC) of the receiver operating characteristic (ROC) curve were compared across the models to evaluate their performance and identify the optimized model for predicting in-ICU mortality of septic patients. In order to compensate for the lack of machine learning algorithms in variable interpretation, we applied SHAP to explain the importance of each feature in the best model.

All statistical analyses were conducted using R programming language (version 4.3.3, the R Foundation, Vienna, Austria) and R Studio software (version 2023.12.1.402, R Studio, PBC, Boston, MA), with R packages including “tidyverse”, “mice”, “rpart”, “dplyr”, “randomForest”, “e1071”, “XGB”, “gbm”, “nnet”, and “shapr”. Statistical significance was decided at a two-sided p -value less than 0.05.

Results

Baseline characteristics

This study included 9288 study participants and 43 variables, covering basic demographic information, laboratory indicators, and comorbidities. Baseline clinical characteristics included ICU ward type, number of days in ICU, pulse, heart rate, temperature, systolic blood pressure, diastolic blood pressure, hemoglobin (Hb), white blood cell (WBC) count, platelet (PLT) count, CCI, SOFA score, SAPSII, etc. Additionally, comorbidities such as diabetes, liver disease, chronic obstructive pulmonary disease, etc., were also incorporated.

Among the 9288 patients included in the study, 5069 patients died in the ICU, with a mortality rate of 54.58%. The CPS group comprised 6622 individuals, among whom 3655 died in the ICU, resulting in a mortality rate of 55.19%. The CNS group consisted of 2666 individuals, with 1414 deaths in the ICU, yielding a mortality rate of 53.04%.

The majority of participants were in the age group of 45 to 79 years (65.20%). There was a slightly different proportion between males and females in the study population, with 57.55% being male and 42.45% female. The majority of patients were of white ethnicity. Chronic lung disease was the most common comorbidity, affecting 1866 out of 9288 (30.60%) septic patients. This was followed by hypertension and liver disease, present in 23.41% and 8.5%, respectively. Selected

features of included patients are shown in Table I. This table compares the baseline characteristics of CPS and CNS sepsis patient subgroups in the ICU. Statistically significant differences were observed in age distribution, ethnicity, and ICU ward type ($p < 0.05$). Severity scores (SOFA, SAPSII) were

Table I. Comparison of baseline characteristics between the two subgroups in the ICU

Variables	CPS group (N = 6622)	CNS group (N = 2666)	Z/ χ^2	P-value
Age	65.48 \pm 16.34	65.52 \pm 16.63	73.38	0.53
Age group/(n, %)			11.47	0.04
18–25	98 (1.48)	56 (2.1)		
26–44	656 (9.91)	277 (10.39)		
45–64	2307 (34.84)	862 (32.33)		
65–79	2042 (30.84)	845 (31.7)		
80–85	740 (11.17)	328 (12.3)		
> 85	779 (11.76)	298 (11.18)		
Male/(n, %)	3802 (57.41)	1543 (57.88)	0.15	0.70
Ethnicity/(n, %)			26.15	< 0.01
Asian	217 (3.28)	52 (1.95)		
Black	739 (11.16)	277 (10.39)		
White	4379 (66.13)	1744 (65.42)		
Other	1287 (19.44)	593 (22.24)		
ICU ward type/(n, %)			212.31	< 0.01
CCU	518 (7.82)	335 (12.57)		
CVICU	469 (7.08)	310 (11.63)		
MICU	2115 (31.94)	629 (23.59)		
MICU/SICU	1781 (26.9)	524 (19.65)		
SICU	940 (14.20)	429 (16.09)		
Neuro SICU	129 (1.95)	83 (3.11)		
TSICU	670 (10.12)	356 (13.35)		
Severity score/(Md [IQR])				
SOFA	7.73 (3.72, 11.75)	7.04 (3.29, 10.8)	82.83	< 0.01
SAPSII	42.35 (27.62, 57.07)	40.27 (26.22, 54.31)	80.92	< 0.01
CCI	6.17 (3.26, 9.09)	6.06 (3.09, 9.04)	76.79	0.15
Comorbidity/(n, %)				
Chronic pulmonary disease	2067 (31.21)	775 (29.07)	4.02	0.05
Liver disease	571 (8.62)	218 (8.18)	0.43	0.51
Diabetes	271 (4.09)	97 (3.64)	0.91	0.34
Hypertension	1554 (23.47)	620 (23.26)	0.04	0.85
Treatment measures/(n, %)				
Ventilation	4369 (65.98)	1740 (65.27)	0.40	0.53
Dialysis	494 (7.46)	180 (6.75)	1.33	0.25
Vasopressin	1083 (16.35)	362 (13.58)	10.94	< 0.01
Number of days/(d[IQR])				
Number of days in hospital	17.2 (0.4, 34.01)	16.07 (0.21, 31.93)	78.28	< 0.01
Number of days in ICU	7.4 (–0.95, 15.75)	7.63 (–0.92, 16.17)	74.11	0.22
Outcome/(n, %)				
Deaths	3655 (55.19)	1414 (53.04)	3.48	0.06
90-day deaths	2206 (33.31)	811 (30.42)	7.12	< 0.01

CCU – coronary care unit, CVICU – cardiac vascular intensive care unit, MICU – medical intensive care unit, MICU/SICU – medical/surgical intensive care unit, SICU – surgical intensive care unit, Neuro SICU – neurosurgical intensive care unit, TSICU – trauma surgical intensive care unit, SOFA – Sequential Organ Failure Assessment, SAPSII – Simplified Acute Physiology Score II, CCI – Charlson Comorbidity Index.

generally higher for CPS patients, indicating greater illness severity ($p < 0.01$). CNS patients showed slightly lower mortality rates and were more likely to be in certain ICU wards such as the Coronary Care Unit (CCU) and Trauma and Surgical Intensive Care Unit (TSICU). Both groups had similar comorbidity patterns, but CPS patients required more vasopressin treatment ($p < 0.01$). Lengths of hospital stay were also significantly different between the groups ($p < 0.01$).

Variable selection

The final selection results are illustrated in the Figure below.

Based on the selection results, we considered the following 14 variables for machine learning modeling analysis with the whole study sample, including number of days in hospital, Hb, hematocrit (Hct), temperature, urine output, GCS, CCI, RBC count, SAPSII, alkaline phosphatase (ALP), blood urea nitrogen (BUN), liver disease, vasopressin, and age group.

In addition, we found that BUN and liver disease were not significant, and they were excluded from the analysis of the CPS group, while mean blood pressure (MBP) and ICU ward type were included due to their significance ($p < 0.05$) in screening stage. Four variables – number of days in hospital, urine output, SAPSII, and liver disease – were not significant in the CNS group and were excluded from its analysis, whereas blood lactic acid (LaC) count and ICU ward type were meaningful variables and included in the modeling analysis of the CNS group.

Model development and validation

Six machine learning models were constructed to predict the risk of in-ICU mortality among

septic patients. We divided datasets into training and test sets in a ratio of 7 : 3, and the prediction results of the six machine learning algorithms based on test sets are shown in Table II. The table describes the prediction performance of the six machine learning models (DT, RF, NB, XGB, GBDT, and NNET) in predicting the in-ICU mortality rate of patients with sepsis, including the CPS and CNS subgroups, which mainly includes accuracy, sensitivity, specificity, and AUC. (The prediction results of the six machine learning models on training sets are shown in Supplementary Table SII.)

For septic patients (the whole study sample), the XGB model achieved the highest AUC of 0.782, while the AUCs of the remaining five models were lower, with the AUC of the DT model being the lowest at 0.6741. Regarding predictive accuracy, the RF, XGB, GBDT, and NNET models all achieved around 70%, while the DT and the NB achieved the lowest accuracy of around 67%. In terms of sensitivity and specificity, XGB, GBDT, and NNET performed better.

Considering the above four evaluation indicators collectively, the XGB algorithm demonstrated the best predictive performance, followed by the NNET model. Additionally, under the same evaluation indicators, the best prediction model for both the CPS and CNS groups was GBDT. XGB outperformed other models in the overall sample, mainly due to its strong generalization ability and adaptability to large datasets. The superiority of GBDT in the two subgroups of CPS and CNS, on the other hand, may stem from its flexibility and robustness in dealing with small samples and data with large feature differences. The ROC curves for the established models are shown in Figure 2.

Table II. Predictive results of six machine learning models based on test sets for septic patients who died in the ICU

Group	Indicators	Model					
		DT	RF	NB	XGB	GBDT	NNET
All patients	Accuracy	0.6701	0.7003	0.6701	0.7136	0.7042	0.6978
	Sensitivity	0.6605	0.6888	0.7062	0.7127	0.7056	0.7054
	Specificity	0.6887	0.7207	0.6324	0.7149	0.7022	0.6872
	AUC area	0.6741	0.7695	0.7372	0.782	0.7816	0.7768
CPS group	Accuracy	0.6934	0.7135	0.6964	0.708	0.7185	0.7044
	Sensitivity	0.6845	0.7136	0.7243	0.7142	0.7281	0.7226
	Specificity	0.7103	0.7134	0.6626	0.6986	0.7049	0.6803
	AUC area	0.6869	0.7775	0.7516	0.7756	0.7779	0.7751
CNS group	Accuracy	0.6683	0.6884	0.6821	0.6871	0.7109	0.6583
	Sensitivity	0.6616	0.7201	0.7709	0.7223	0.7403	0.7106
	Specificity	0.6872	0.6435	0.6051	0.6391	0.6697	0.5967
	AUC area	0.661	0.7558	0.754	0.7439	0.7588	0.7223

DT – decision tree, RF – random forest, NB – naive bayes, XGB – Extreme Gradient Boosting, GBDT – Gradient Boosted Decision Trees, NNET – Neural Network, AUC – area under the curve.

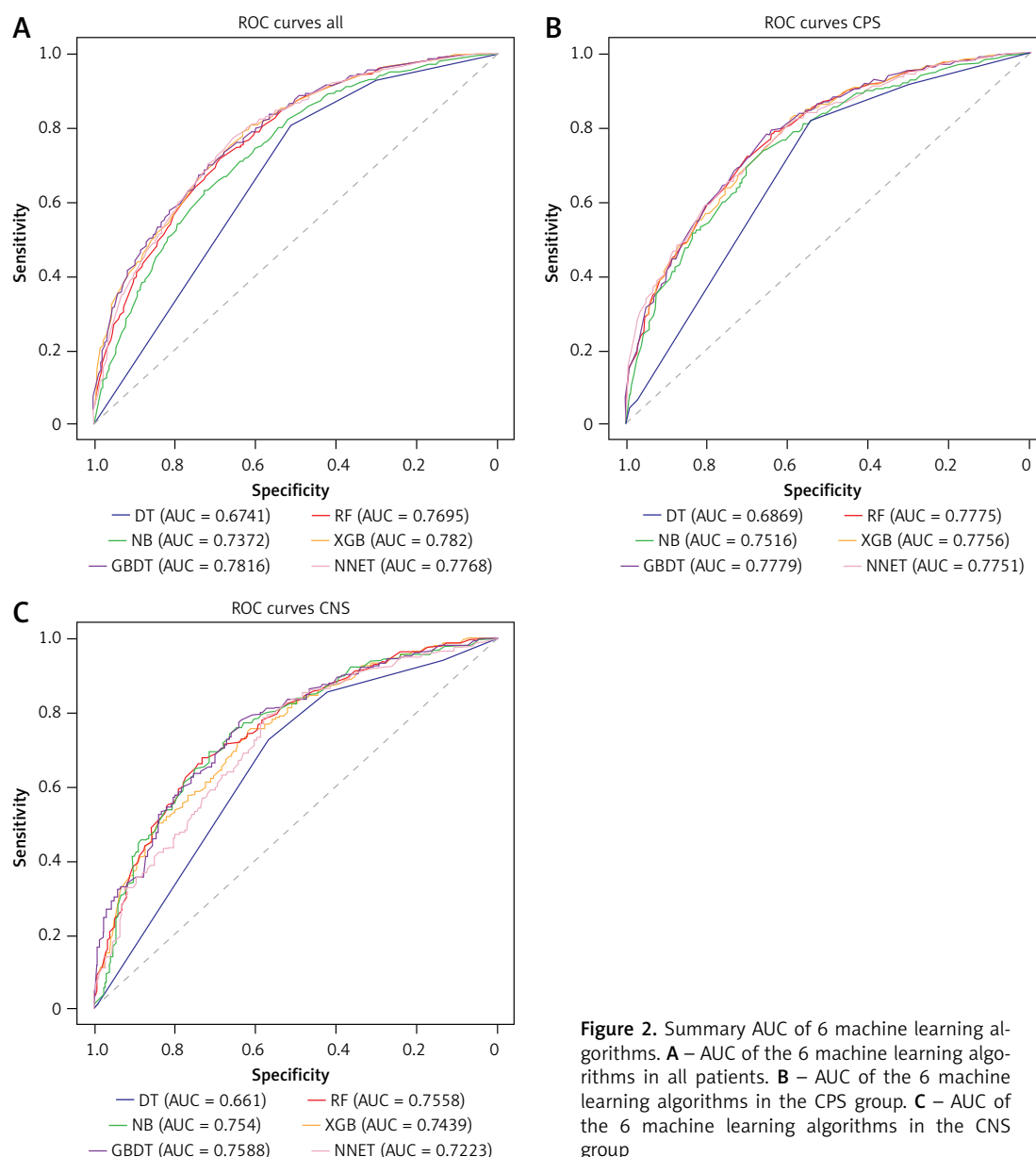


Figure 2. Summary AUC of 6 machine learning algorithms. **A** – AUC of the 6 machine learning algorithms in all patients. **B** – AUC of the 6 machine learning algorithms in the CPS group. **C** – AUC of the 6 machine learning algorithms in the CNS group

Significance of features

CCI, SAPSII, BUN, liver disease, vasopressin, and age were associated with a higher risk of in-ICU death. As their levels increased, the patient's risk of death increased. Number of days in hospital, Hb count, temperature, urine output, GCS, and RBC count were associated with lower in-ICU mortality. Patients in Cardiac Vascular Intensive Care Unit (CVICU), Surgical Intensive Care Unit (SICU), and Trauma SICU (TSICU) had a higher risk of in-ICU mortality compared to the Coronary Care Unit (CCU). Risk factors that were the same between the CPS and CNS subgroups were CCI, age, GCS, bilirubin total, temperature, Hb, RBC count, ICU ward type, hematocrit, and vasopressin. In terms of differences between the two subgroups, days of hospitalization, urine output, SAPSII, and MBP were

risk factors specific to the CPS group that may lead to an increased risk of death in the ICU, whereas in the CNS group, higher levels of BUN and LaC were specific risk factors (Figures 3 and 4).

Discussion

In this study, the clinical measurements of sepsis patients were extracted from the MIMIC-IV database. Then, we compared the baseline features and clinical characteristics of patients in the CPS and the CNS subgroups. We identified three different combinations of factors that may be associated with in-ICU death of sepsis patients. Based on six different machine learning algorithms, prediction models for in-ICU death were established for general patients with sepsis, as well as the CPS and the CNS groups. The optimized prediction

model was selected by comprehensively considering the evaluation metrics, including AUC, prediction accuracy, sensitivity and specificity. The XGB model performed best for general patients with sepsis, while the GBDT model performed best for the CPS and the CNS groups. Based on the interpretable learning method, we explored the importance of the identified relevant factors for the prognosis of sepsis patients.

Despite the many challenges of studying patients with CNS, exploring the characteristics of patients with CNS can further reveal their prognostic features and influencing factors, which can contribute to treatment efficacy and improve prognosis [7]. By identifying specific biomarkers or clinical features, it is also possible to provide useful information for the development of treatment strategies for patients with CPS and CNS

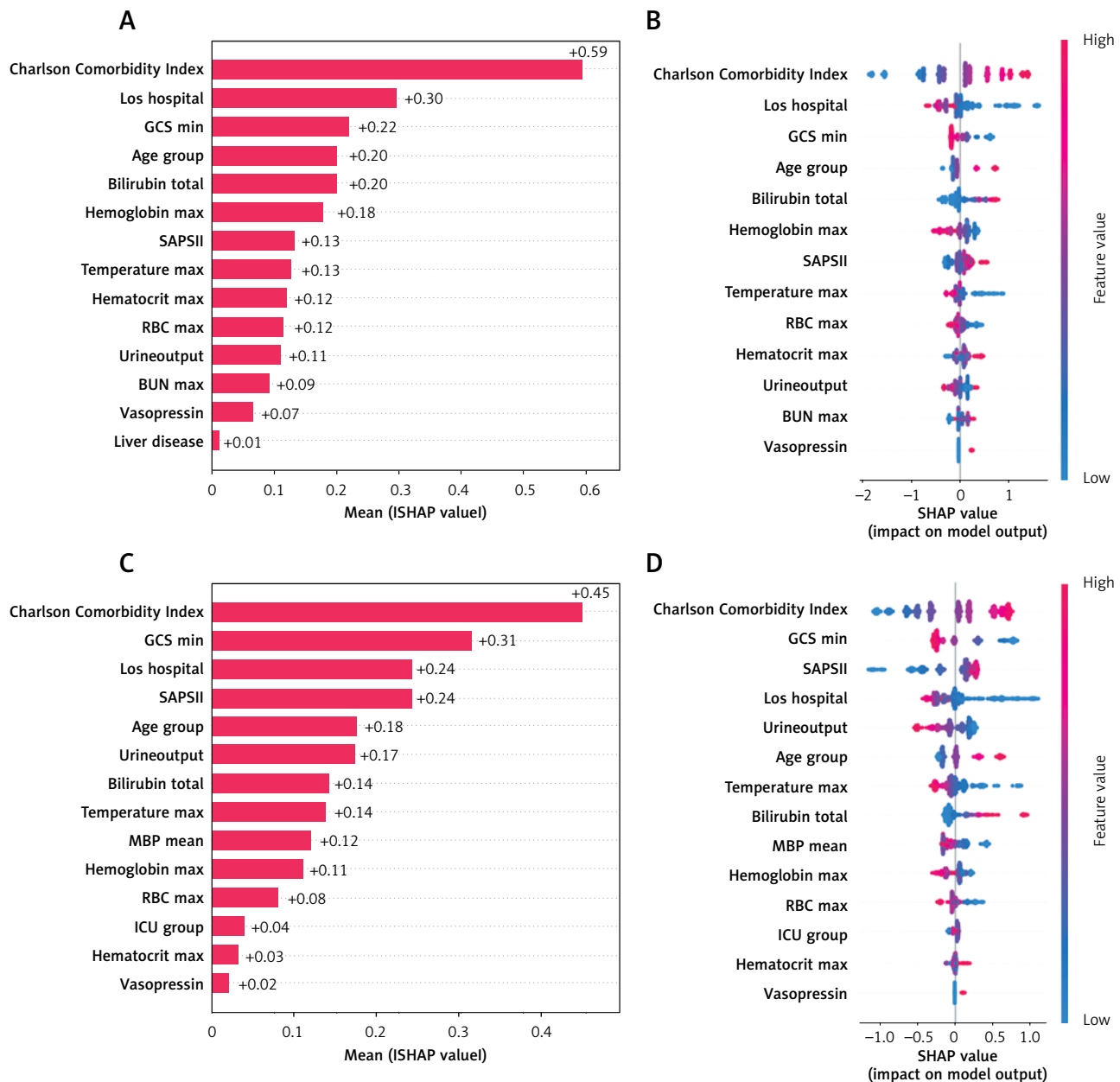


Figure 3. Summary of SHAP for the XGB model for all patients, the GBDT model for the CPS group, and the GBDT model for the CNS group. **A** – Feature importance from the XGB model of all patients. Feature importance is represented by the absolute mean of the SHAP values for each feature. **B** – The higher the SHAP feature value, the higher the risk of death in the ICU for patients with sepsis. A point created in the model represents a characteristic attribute value for one patient, so that there is a point on each characteristic line for each patient. The points are colored according to the characteristic value of the corresponding patient and vertically cumulated to describe the density. Red color indicates higher eigenvalues and blue color indicates lower eigenvalues. **C** – Feature importance from the GBDT model of the CPS group. Feature importance is represented by the absolute mean of the SHAP values for each feature. **D** – The higher the SHAP feature value, the greater the risk of death in the ICU for CPS patients

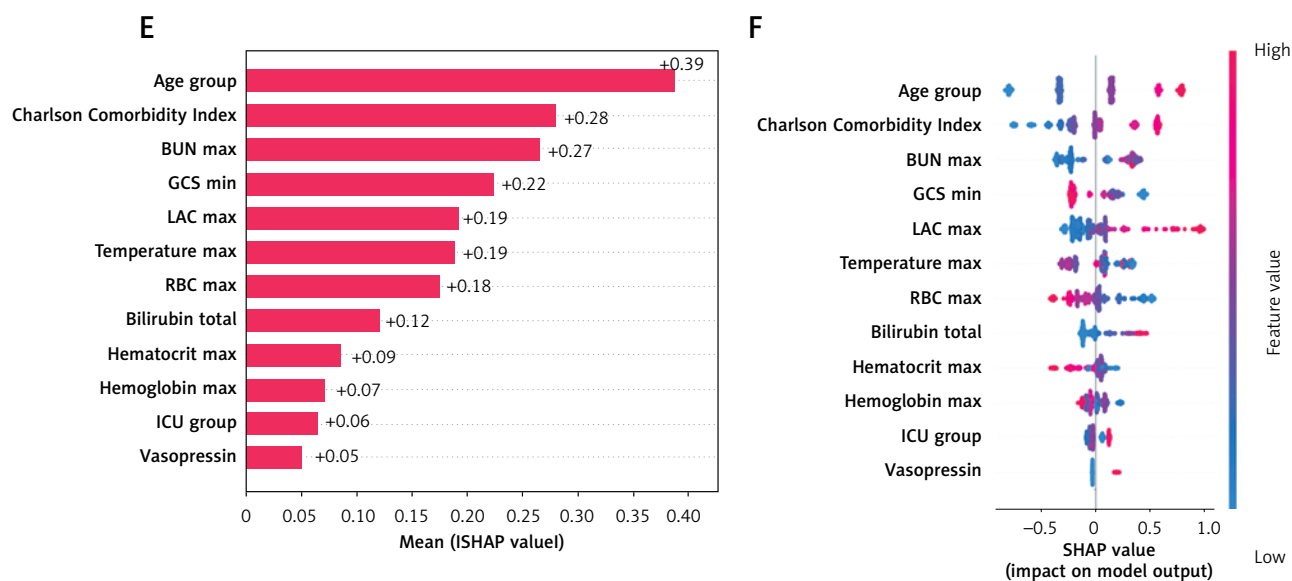


Figure 3. Cont. **E** – Feature importance from the GBDT model of the CNS group. Feature importance is represented by the absolute mean of the SHAP values for each feature. **F** – The higher the SHAP feature value, the greater the risk of death in the ICU for CNS patients

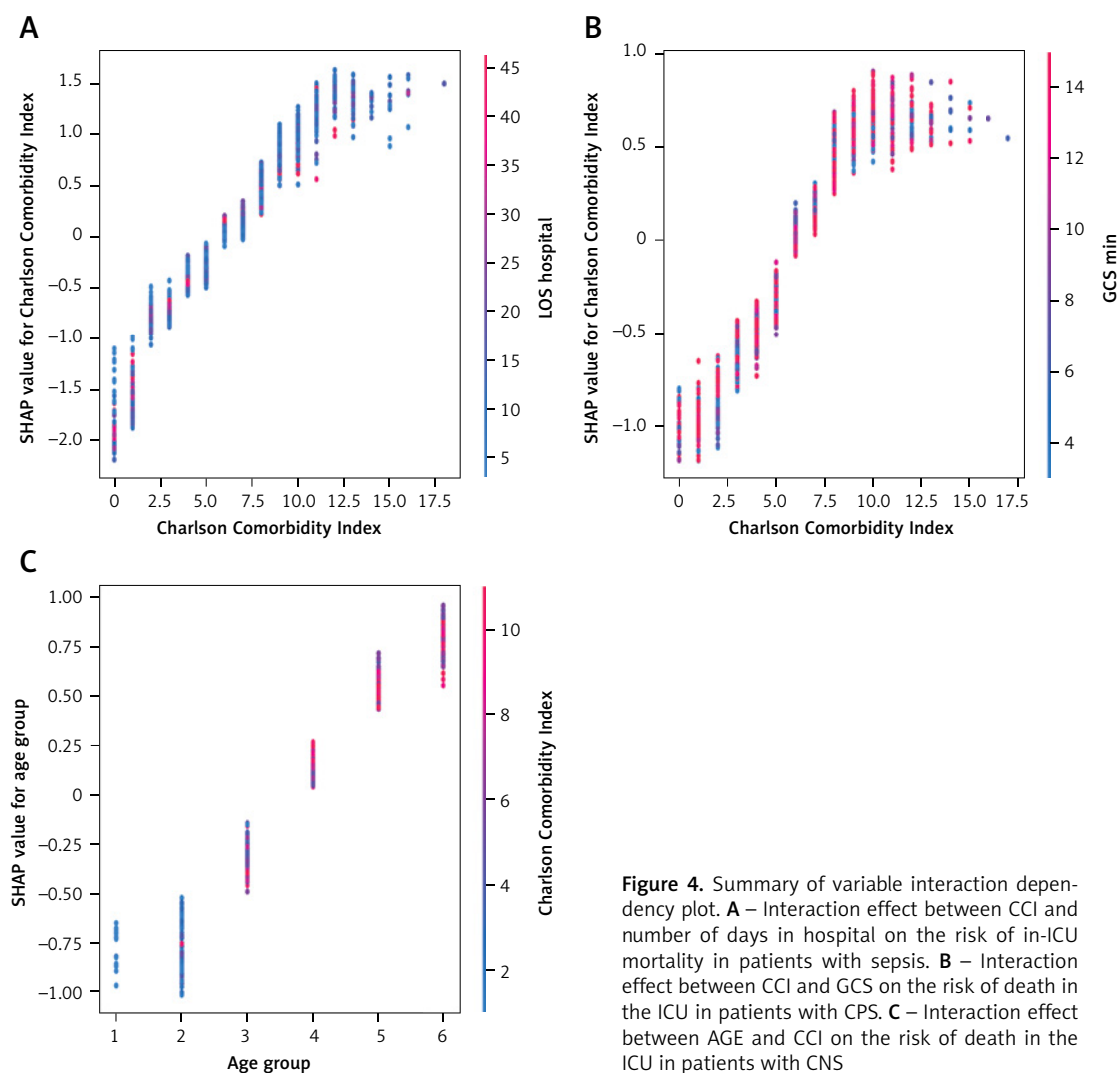


Figure 4. Summary of variable interaction dependency plot. **A** – Interaction effect between CCI and number of days in hospital on the risk of in-ICU mortality in patients with sepsis. **B** – Interaction effect between CCI and GCS on the risk of death in the ICU in patients with CPS. **C** – Interaction effect between AGE and CCI on the risk of death in the ICU in patients with CNS

[5]. The machine learning models developed offer a valuable tool for ICU clinicians, helping to identify high-risk patients early and enabling timely, targeted interventions [13]. Integrating these models into electronic health record (EHR) systems could facilitate real-time risk assessment and personalized care [21]. Future steps include prospective validation in independent ICU cohorts and the incorporation of dynamic data (e.g., real-time vital signs) to enhance model accuracy. Ultimately, these models could improve patient outcomes by guiding interventions and resource allocation, potentially reducing in-ICU mortality.

Among the clinical features selected, low levels of hospital days, Hb, urine output, temperature, and GCS were found to be associated with increased risk of in-ICU death. Higher levels of age, Hct, CCI, SAPSII, TBil, BUN, the presence of liver disease, and use of vasopressin were also associated with a higher in-ICU death risk for sepsis patients.

Our analysis revealed little difference in comorbidities between the patients with CPS and CNS, but patients with CPS experienced more vasopressor support, which may be related to the severely impaired endothelial function of patients with CPS. It has been shown [22] that sepsis is an abnormal host response to infection, which is characterized by a variety of systemic disturbances, including increased vascular permeability and vasodilatation. Therapeutically, in addition to the use of antibiotics [23], early fluid resuscitation is also critical [24]. Su *et al.* reported that a large positive fluid balance worsens the condition of patients with hypoxemia [25]. However, not all patients respond effectively to fluid resuscitation, and its effects usually last for a shorter period [26], thus requiring the use of vasoconstrictive drugs to improve organ perfusion [27]. In addition, septic patients admitted to different types of ICU wards differ in their risk of in-ICU death, and a study by Ohbe *et al.* found a difference in case-fatality rates between patients in the ICU and those in the general ward [28], but no study has yet explored the association between the type of ICU ward and the risk of in-ICU death in septic patients.

Our study further compared the similarities and differences in identified influential factors and established models for patients with CPS and CNS. Risk factors that were the same between the two subgroups were CCI, age, GCS, bilirubin total, temperature, Hb, RBC count, ICU ward type, hematocrit, and vasopressin. This may be related to disease risk factors in sepsis patients, which did not seem to differ between the CPS and CNS groups. Nejtek *et al.* found that a higher CCI had a significant effect on mortality [29]. For differences between the two subgroups, number of days

in hospital, urine output, SAPSII, and mean MAP were important factors related to increased risk of in-ICU death in the CPS group. However, in the CNS group, these variables were not included in the final model. In the CNS group, higher BUN and LaC levels were significant risk factors related to increased risk of in-ICU death, whereas they were not included in the final model in the CPS group. Sigakis *et al.* observed that blood transfusion within the first 24 h of onset was significantly associated with mortality in patients with CNS, but not in patients with CPS [30]. These differences may reflect variations in the type of pathogen infection and severity of the disease between the two groups, requiring targeted prevention and treatment measures for each group. Therefore, based on these differences, we can further explore the potential pathological and physiological similarities and differences between patients with CPS and CNS in order to achieve meaningful results in the prevention or treatment of sepsis in the future.

In terms of model performance, the XGB model performed best for predicting death in the ICU for all patients. For patients in the CPS and CNS subgroups, the GBDT model performed best. These findings are similar to those of Yang *et al.*, who reported that XGB and RF models performed optimally for predicting sepsis onset [17]. However, the choice of model also needs to be based on practical clinical scenarios and needs to be weighed against other considerations.

In this study, the SHAP approach was also used to enhance the interpretation of the model established and to explore the main risk factors affecting sepsis patients' mortality in the ICU. The results showed that some common risk factors, such as higher TBil, renal insufficiency (reflected in higher BUN levels), higher SAPSII, older age, and treatment with vasopressin significantly increased the risk of in-ICU death for the whole study sample and in two subgroups. This is consistent with the results of previous studies [1, 31, 32]. It has been shown that the prevalence of sepsis is increasing in the elderly population and age is an independent risk factor for death [31]. A study by Kübler *et al.* revealed that about 89% of septic patients admitted to the ICU in critical condition had dysfunction in three or more organs [32]. Organ dysfunction can be measured by the SOFA score, where an increase of 2 or more points implies an in-hospital mortality rate of more than 10% [1].

We also found that lower Hb levels, urine output, and GCS were associated with a lower risk of in-ICU death. These findings are also supported by previous studies [32–35]. Qi *et al.* found that when a patient's Hb level falls below 80 g/l within 48 h of admission to the ICU, it will likely increase

the risk of death in patients with sepsis [33]. Low urine output is an important indicator for detecting acute kidney injury [36], and also suggests an increased risk of death [34], which is often associated with renal dysfunction in patients with sepsis [32]. GCS is a rating scale that provides an objective assessment of patient consciousness [35], and a low GCS may be a simple and reliable predictor suggesting the need for initial resuscitation in patients with sepsis [37]. These findings are in line with the conclusions of previous studies, suggesting that, to improve the prognosis of septic patients, we may need to closely monitor the liver and kidney functions, SOFA score and other indicators of infected patients, actively correct the anemia, and maintain the patient's urine output and mental status, so as to allow the patient to obtain a good therapeutic prognosis.

The SHAP approach applied provide better understanding of the final machine learning model and selected influential factors. A comprehensive analysis of the key factors affecting the prognosis of sepsis patients provided a potentially important guide for future risk assessment and individualized treatment in clinical practice. It also provides the quantitative evidence necessary for better understanding of the prognosis of septic patients.

Limitations of this study include the following: (1) Despite our rigorous screening of potential factors, we may not have been able to exclude all possible confounders due to the retrospective nature of this study, and we were unable to consider some important confounders such as antimicrobial use and timing of interventions due to lack of information about these factors. (2) Our analyses were based on retrospective data on the history of the first ICU admission, and potential bias in sample selection may not be totally avoided. (3) The aim of our study was to explore the major influences on bacterial sepsis. The inclusion of other pathogens (e.g., fungi and viruses) may greatly increase the complexity of the data and make it difficult to accurately differentiate the specific characteristics of bacterial infections. Future studies could take fungal and viral infections into account to fully assess the impact of multiple pathogens. (4) Our study has only established a preliminary prediction model, and further prospective validation and optimization are needed before application in practice.

In conclusion, this study found no significant difference in in-ICU mortality between CPS and CNS, but it identified key risk factors for in-ICU mortality in sepsis patients, revealing significant differences between CPS and CNS. These findings can inform personalized treatment strategies, particularly for CNS patients, who may require distinct management approaches.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found at <https://physionet.org/content/mimic-iv-ed/2.2/>.

Funding

This work was supported by the National Social Science Fund of China (21CTJ009).

Ethical approval

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

References

- Schlapbach LJ, Watson RS, Sorce LR, et al. International consensus criteria for pediatric sepsis and septic shock. *JAMA* 2024; 331: 665-74.
- Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet* 2018; 392: 75-87.
- Prescott HC, Angus DC. Enhancing recovery from sepsis: a review. *JAMA* 2018; 319: 62-75.
- Oshima T, Kodama Y, Takahashi W, et al. Empiric antibiotic therapy for severe sepsis and septic shock. *Surg Infect* 2016; 17: 210-6.
- Chun K, Syndergaard C, Damas C, et al. Sepsis pathogen identification. *J Lab Autom* 2015; 20: 539-61.
- Gatica S, Fuentes B, Rivera-Asín E, et al. Novel evidence on sepsis-inducing pathogens: from laboratory to bedside. *Front Microbiol* 2023; 14: 1198200.
- Ou Q, Cai G, Zhou Y, Wen M. [Analysis of clinical features and risk factors for mortality in patients with culture-negative sepsis: a single-center retrospective cohort study based on MIMIC-IV]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 2021; 33: 1181-6.
- Kim JS, Kim YJ, Kim WY. Characteristics and clinical outcomes of culture-negative and culture-positive septic shock: a single-center retrospective cohort study. *Crit Care* 2021; 25: 11.
- Hazwani TR, Kazzaz YM, Alsugheir S, et al. Association between culture-negative versus culture-positive sepsis and outcomes of patients admitted to the pediatric intensive care unit. *Cureus* 2020; 12: e9981.
- Baker AH, Leland SB, Freiman E, Herigon JC, Eisenberg MA. Characteristics and outcomes of culture-positive and culture-negative pediatric sepsis. *J Pediatr* 2023; 263: 113718.
- Phua J, Ngerng W, See K, et al. Characteristics and outcomes of culture-negative versus culture-positive severe sepsis. *Crit Care* 2013; 17: R202.
- Paul M, Shani V, Muchtar E, Kariv G, Robenshtok E, Leibovici L. Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis. *Antimicrob Agents Chemother* 2010; 54: 4851-63.
- 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; 46: 547-53.
- Persson I, Östling A, Arlbrandt M, Söderberg J, Bece-das D. A machine learning sepsis prediction algorithm

- for intended intensive care unit use (NAVOY Sepsis): proof-of-concept study. *JMIR Form Res* 2021; 5: e28000.
15. Yao RQ, Jin X, Wang GW, et al. A machine learning-based prediction of hospital mortality in patients with postoperative sepsis. *Front Med* 2020; 7: 445.
16. Zhou H, Liu L, Zhao Q, et al. Machine learning for the prediction of all-cause mortality in patients with sepsis-associated acute kidney injury during hospitalization. *Front Immunol* 2023; 14: 1140755.
17. Yang Z, Cui X, Song Z. Predicting sepsis onset in ICU using machine learning models: a systematic review and meta-analysis. *BMC Infect Dis* 2023; 23: 635.
18. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset [published correction appears in *Sci Data*. 2023 Jan 16;10(1):31] [published correction appears in *Sci Data*. 2023 Apr 18;10(1):219]. *Sci Data* 2023; 10: 1.
19. Chen S, Hu W, Yang Y, et al. Predicting six-month re-admission risk in heart failure patients using multiple machine learning methods: a study based on the chinese heart failure population database. *J Clin Med* 2023; 12: 870.
20. Münch MM, Peeters CFW, Van Der Vaart AW, Van De Wiel MA. Adaptive group-regularized logistic elastic net regression. *Biostatistics* 2021; 22: 723-37.
21. Bloom BM, Pott J, Thomas S, Gaunt DR, Hughes TC. Usability of electronic health record systems in UK EDs. *Emerg Med J* 2021; 38: 410-5.
22. Ince C, Mayeux PR, Nguyen T, et al. The endothelium in sepsis. *Shock* 2016; 45: 259-70.
23. Manan MM, Ibrahim NA, Aziz NA, Zulkifly HH, Al-Worafi YM, Long CM. Empirical use of antibiotic therapy in the prevention of early onset sepsis in neonates: a pilot study. *Arch Med Sci* 2016; 12: 603-13.
24. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Crit Care Med* 2021; 49: e1063-143.
25. Su L, Liu S, Yang Y, et al. Positive fluid balance and poor outcomes after initial intensive care unit admission in sepsis resuscitation: a retrospective study. *Arch Med Sci* 2024; 20: 464-75.
26. Aya HD, Ster IC, Fletcher N, Grounds RM, Rhodes A, Cecconi M. Pharmacodynamic analysis of a fluid challenge. *Crit Care Med* 2016; 44: 880-91.
27. Ibarra-Estrada M, Kattan E, Aguilera-González P, et al. Early adjunctive methylene blue in patients with septic shock: a randomized controlled trial. *Crit Care* 2023; 27: 110.
28. Ohbe H, Sasabuchi Y, Doi K, Matsui H, Yasunaga H. Association between levels of intensive care and in-hospital mortality in patients hospitalized for sepsis stratified by sequential organ failure assessment scores. *Crit Care Med* 2023; 51: 1138-47.
29. Nejtek T, Müller M, Moravec M, Průcha M, Zazula R. Bacteremia in patients with sepsis in the ICU: does it make a difference? *Microorganisms* 2023; 11: 2357.
30. Sigakis MJG, Jewell E, Maile MD, Cinti SK, Bateman BT, Engoren M. Culture-negative and culture-positive sepsis: a comparison of characteristics and outcomes. *Anesth Analg* 2019; 129: 1300-9.
31. Martin GS, Mannino DM, Moss M. The effect of age on the development and outcome of adult sepsis. *Crit Care Med* 2006; 34: 15-21.
32. Kübler A, Adamik B, Durek G, et al. Results of the severe sepsis registry in intensive care units in Poland from 2003-2009. *Anaesthesiol Intensive Ther* 2015; 47: 7-13.
33. Qi D, Peng M. Early hemoglobin status as a predictor of long-term mortality for sepsis patients in intensive care units. *Shock* 2021; 55: 215-23.
34. Priyanka P, Zarbock A, Izawa J, Gleason TG, Renfurm RW, Kellum JA. The impact of acute kidney injury by serum creatinine or urine output criteria on major adverse kidney events in cardiac surgery patients. *J Thorac Cardiovasc Surg* 2021; 162: 143-51.e7.
35. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974; 2: 81-4.
36. Kaddourah A, Basu RK, Bagshaw SM, Goldstein SL; AWARE Investigators. Epidemiology of acute kidney injury in critically ill children and young adults. *N Engl J Med* 2017; 376: 11-20.
37. Matsuda W, Kimura A, Uemura T. The reverse shock index multiplied by the Glasgow Coma Scale score can predict the need for initial resuscitation in patients suspected of sepsis. *Glob Health Med* 2023; 5: 223-8.