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

Abstract

Introduction
To identify the key risk factors influencing in-ICU mortality of patients with sepsis and develop prognosis prediction models for patients with 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, including DT, RF, NB, XGB, GBDT and NNET, to predict in-ICU death for all study samples, CPS and CNS subgroups, respectively. Model performance were assessed using AUC, accuracy, sensitivity, and specificity. 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. Main Significant influential factors identified included Charlson comorbidity index (CCI), number of days in hospital, Glasgow Coma Scale (GCS), older age, 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 highlights 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.

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

Sitong Liu [1], Lingmin Gong [1], Weiwei Hu [1], Jiaxin Cai [1], Yuhui Yang [1], Shiyu Chen [1], Baibing Mi [1], Yaling Zhao [1], Leilei Pei [1], Fangyao Chen [1,2*]

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

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

*: correspondence to: Fangyao Chen (chenfy@xjtu.edu.cn)

Abstract

Aim: To identify the key risk factors influencing in-ICU mortality of patients with sepsis and develop prognosis prediction models for patients with culture-positive sepsis (CPS) and culture-negative sepsis (CNS) patients.

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, including DT, RF, NB, XGB, GBDT and NNET, to predict in-ICU death for all study samples, CPS and CNS subgroups, respectively. Model performance were 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. Main Significant influential factors identified included Charlson comorbidity index (CCI), number of days in hospital, Glasgow Coma Scale (GCS), older age, 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).

Conclusion: This study identified key risk factors for in-ICU death in patients with sepsis and highlights 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.

1. 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 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 CPS and CNS sepsis, based on pathogen detection within 24 hours 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 does not indicate a significant link 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, Alexandra et al. found that patients in the CPS and CNS groups exhibited similar symptoms upon admission and received comparable treatments[10].

Among existing studies, sepsis research predominantly focuses 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 like 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 outcomes【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 like 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 aims to identify in-ICU mortality risk factors, compare the prognostic differences between CPS and CNS patients, and create predictive models using machine learning techniques.

2 Methods and Materials

2.1 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) database. 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 existing literature [18].

2.2 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 criterion.

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 hours; (4) Microbial cultures performed within 24 hours before and after ICU admission.

Exclusion Criteria: (1) Patients diagnosed with sepsis after 24 hours 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 cases (71.30%) in the CPS group and 2666 cases (28.70%) in the CNS group. (Figure 1)
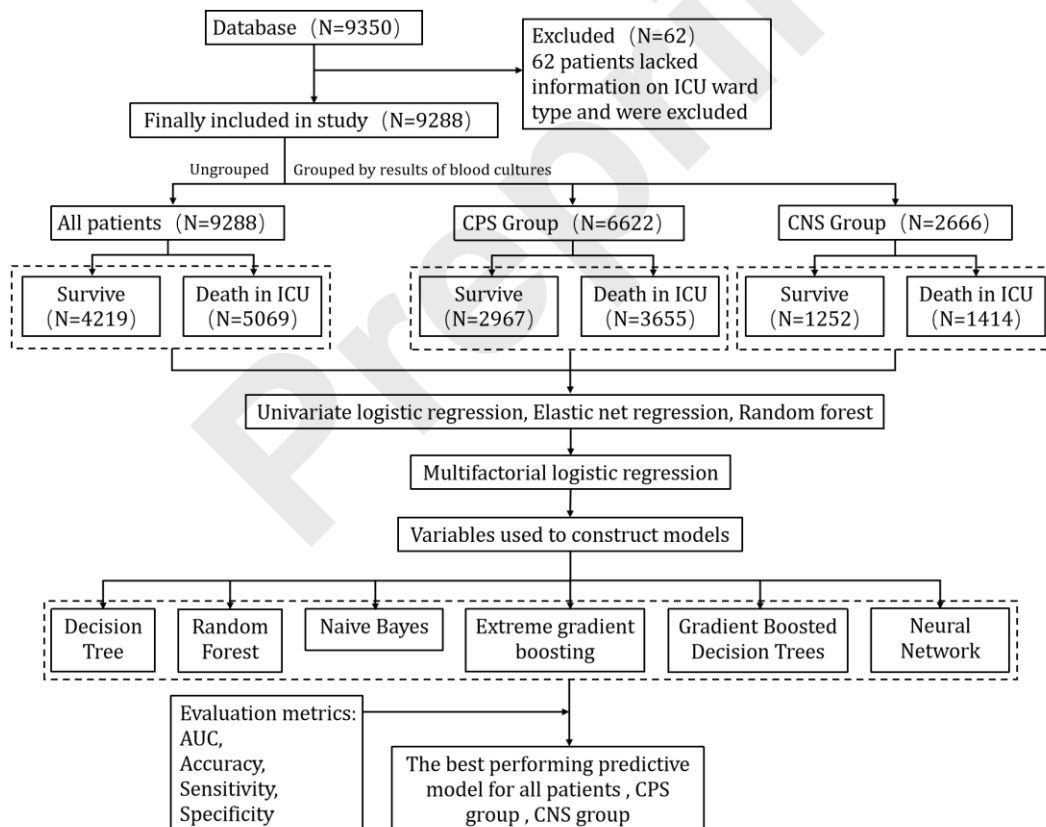

Figure1. Study design. AUC : area under the curve.

2.3 Study Outcome

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

septic patients.

2.4 Data Pre-Processing

Variables with ≥ 30% missingness were excluded from the analysis due to high missingness. Multiple imputation was performed on the remaining missing variables. To further validate the robustness of the interpolation method, we use 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 train the machine learning model on each of the interpolation datasets separately to obtain the prediction results. And then integrate the interpolated datasets using the pool function to achieve the final analyzing dataset. 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 that had received vasopressin or ventilation or dialysis as 1. (The filtered variables and variable missingness rates are shown in Supplementary Table S1)

2.5 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.

2.6 Statistical Analysis

Descriptive analysis was performed on all included patients. Continuous variables conforming to a normal distribution were described using means ± 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 including 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 use the 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.

3 Results

3.1 Baseline Characteristics

This study included 9288 study participants and 43 variables, covering basic demographic informations, 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 ICU, with a mortality rate of 54.58%. The CPS group comprised 6622 individuals, among whom 3655 died in ICU, resulting in a mortality rate of 55.19%. The CNS group consisted of 2666 individuals, with 1414 deaths in ICU, yielding a mortality rate of 53.04%.

The majority of participants were in the age group of 45 to 79 years, accounting for 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, with 1866 out of 9288 septic patients affected, representing 30.60%. This was followed by hypertension and liver disease, accounting for 23.41% and 8.5%, respectively. Selected features of included patients are shown in Table 1. This table compares the baseline characteristics of CPS and CNS sepsis patient subgroups in the ICU. Key differences were observed in age distribution, ethnicity, and ICU ward type, with statistically significant variations (p < 0.05). Severity scores (SOFA, SAPSII) were 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 like CCU and 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).

Table 1 Comparison of baseline characteristics between the two subgroups in ICU

Variables	CPS group(N=6622)	CNS group(N=2666)	Z/χ^2	P value
age	65.48±16.34	65.52±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
90day 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 = Neuro Surgical Intensive Care Unit, TSICU = Trauma Surgical Intensive Care Unit, SOFA = Sequential Organ Failure Assessment, SAPSII = simplified acute physiology score Ⅱ, CCI = Commodity Channel Index.

3.2 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), BUN, liver disease, vasopressin and age group.

In addition, we found that BUN and liver disease were not significant and excluded for the analysis

with CPS group, while mbp and ICU ward type were included due to their significance ($P<0.05$) in screening stage. Four variables including number of days in hospital, urine output, SAPSII, and liver disease were not significant in the CNS group and excluded in formal analysis with the CNS group, whereas Blood Lactic Acid (LaC) count and ICU ward type were meaningful variables and considered into the modeling analysis of the CNS group.

3.3 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 2. 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 S2.)

Table2 Predictive results of six machine learning models based on test sets for septic patients die 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.

For septic patients (the whole study sample), it can be learnt that the XGB model achieved the highest AUC of 0.782, while the AUCs of the remaining five models were relatively lower, with the AUC of the DT model being the lowest at 0.6741. Regarding predictive accuracy, the RF, the XGB,

the GBDT and the 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, the XGB, the GBDT and the 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 is GBDT. XGB outperforms 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.
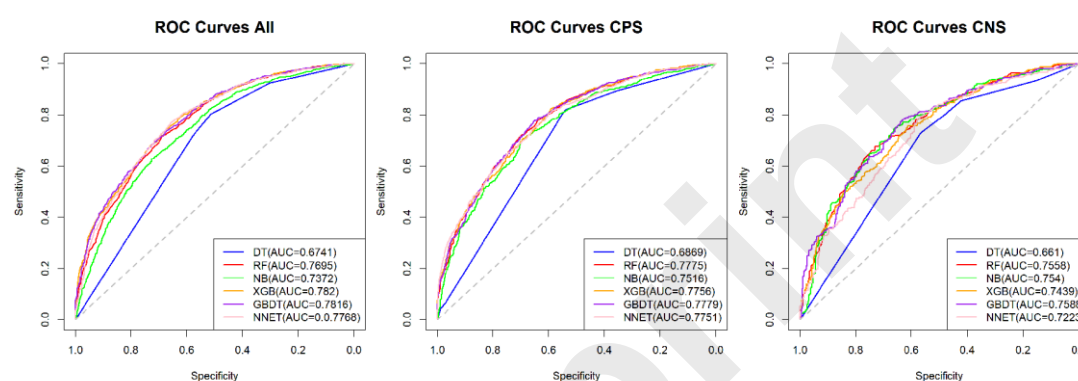

Figure 2. Summary AUC of 6 machine learning algorithms. (A) The AUC of the 6 machine learning algorithms in the All patients. (B)The AUC of the 6 machine learning algorithms in the CPS group. (C)The AUC of the 6 machine learning algorithms in the CNS group.

3.4 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 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.
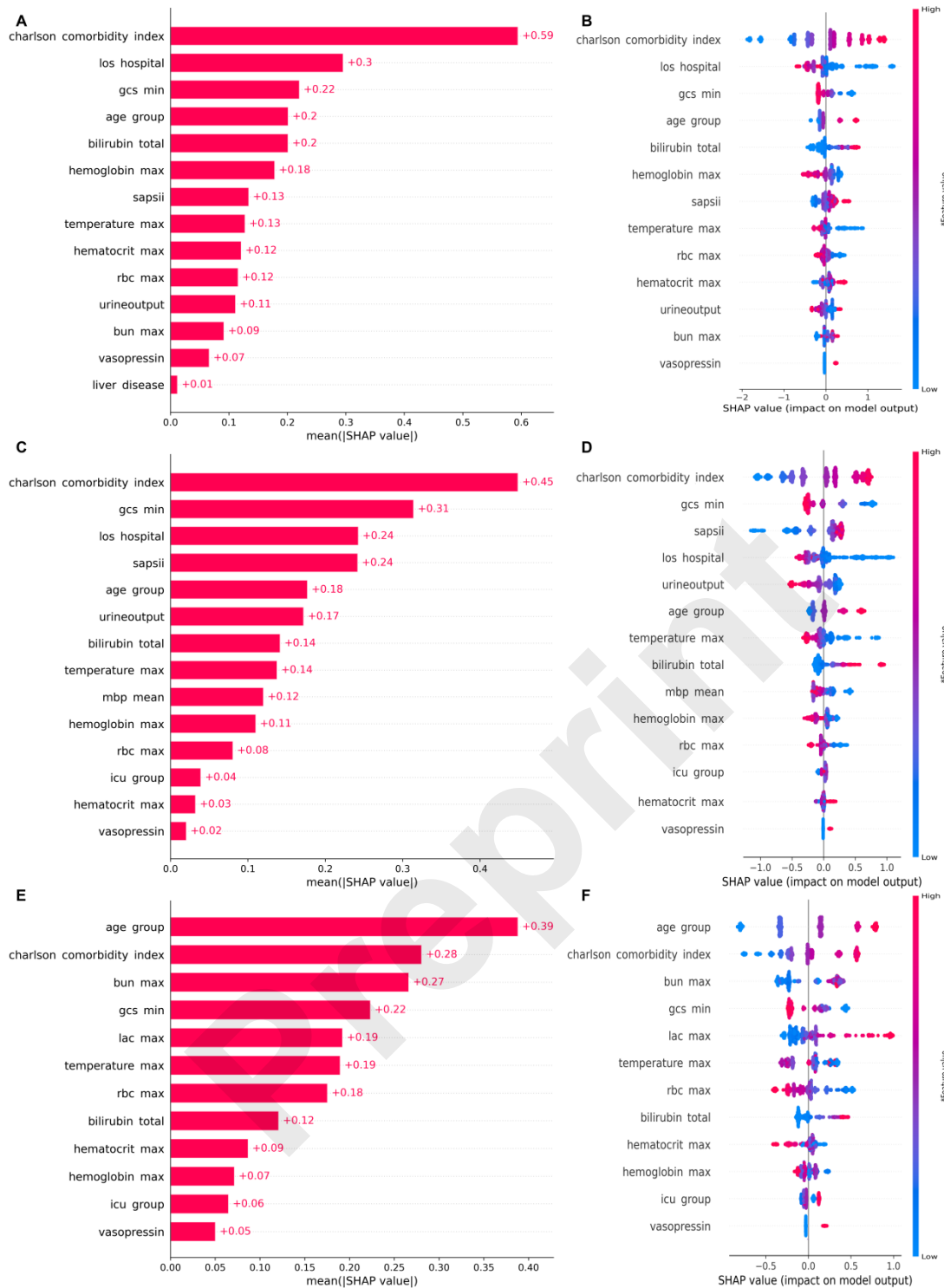
Figure 3. Summary of SHAP for the XGB model for the All patients, the GBDT model for the CPS group, and the GBDT for the CNS group. (A) Feature importance from the XGB model of the All patients and the absolute value of the mean of the SHAP values for each feature is the feature importance distribution. (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 and the absolute value of the mean of the SHAP values for each feature is the feature importance distribution. (D) The higher the SHAP feature value, the greater the risk of death in the ICU for CPS patients. (E) Feature importance from the GBDT model of the CNS group and the absolute value of the mean of the SHAP values for each feature is the feature importance distribution. (F) The higher the SHAP feature value, the greater the risk of death in the ICU for CNS patients.
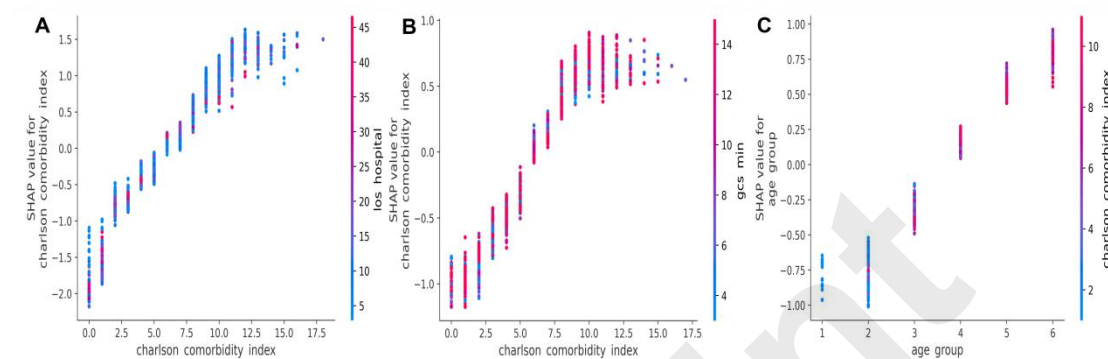

Figure 4. Summary of variable interaction dependency plot. (A) Interaction between CCI and Number of days in hospital on the risk of in-ICU mortality in patients with sepsis. (B)Interaction between CCI and GCS on the risk of death in the ICU in patients with CPS.. (C) Interaction between AGE and CCI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

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 of AUC, prediction accuracy, sensitivity and specificity. The XGB model performed best for general patients with sepsis, the GBDT model performed best for the CPS and the CNS groups. And based on the interpretable learning method, we explored the importance of the identified relevant factors on 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[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, are 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 are also associated with a higher in-ICU death risk for sepsis patients.

Our analysis found 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. showed that the condition of patients with hypoxemic is deteriorated by a large positive fluid balance[25]. However, not all patients respond effectively to fluid resuscitation and its effects usually maintain for a shorter period of time[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 ICU and those in general ward[28], but no study has yet delved into 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, MAP mean 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, a higher BUN and LaC level was a significant risk factor related to increased risk of in-ICU death, whereas it was not included in the final model in the CPS group. Matthew et al. showed that blood transfusion within the first 24 hours of onset was significantly associated with mortality in patients with CNS, but not in patients with CPS[30]. These differences may be related to the existence of differences 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 ICU for all patients. For patients in the CPS and CNS subgroups, the GBDT model performed best. These findings are similar to Yang et al. who suggested that XGB and RF 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, SHAP approach were also used to enhance the interpretation of the model established and to explore the main risk factors affecting sepsis patients' mortality in ICU. The results showed that some common risk factors such as higher TBil, renal insufficiency (reflected in higher BUN levels), higher SAPSII, increased 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 existing 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. found that about 89% of septic patients admitted to ICU in critical condition had dysfunction in three or more organs[32]. Organ dysfunction can be measured by SOFA score, where an increase of 2 or more points in SOFA score 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 hours of admission to 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, in order to promote 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 anaemia, 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. Comprehensively sorting out the key factors affecting the prognosis of sepsis patients provided a potentially important guide for future risk assessment and individualized treatment in the clinical practice. It also provides the quantitative evidence necessary for better understanding of the prognosis of septic patients.

Limitations of this study include: (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 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.

5. Conclusions

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. This data can be found here: https://physionet.org/content/mimic-iv-ed/2.2/.

Ethical approval

Not applicable, because this article does not contain any studies with human or animal subjects.

References:

[1] Schlapbach LJ, Watson RS, Sorce LR, et al. International Consensus Criteria for Pediatric Sepsis and Septic Shock. JAMA. 2024;331(8):665-674. doi:10.1001/jama.2024.0179

[2] Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. Lancet. 2018;392(10141):75-87. doi:10.1016/S0140-6736(18)30696-2

[3] Prescott HC, Angus DC. Enhancing Recovery From Sepsis: A Review. JAMA. 2018;319(1):62-75. doi:10.1001/jama.2017.17687

[4] Oshima T, Kodama Y, Takahashi W, et al. Empiric Antibiotic Therapy for Severe Sepsis and Septic Shock. Surg Infect (Larchmt). 2016;17(2):210-216. doi:10.1089/sur.2014.096

[5] Chun K, Syndergaard C, Damas C, et al. Sepsis Pathogen Identification. J Lab Autom. 2015;20(5):539-561. doi:10.1177/2211068214567345

[6] Gatica S, Fuentes B, Rivera-Asín E, et al. Novel evidence on sepsis-inducing pathogens: from laboratory to bedside. Front Microbiol. 2023;14:1198200. Published 2023 Jun 23. doi:10.3389/fmicb.2023.1198200

[7] Ou Q, Cai G, Zhou Y, Wen M. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue. 2021;33(10):1181-1186. doi:10.3760/cma.j.cn121430-20210525-00780

[8] 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(1):11. Published 2021 Jan 6. doi:10.1186/s13054-020-03421-4

[9] 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(8):e9981. Published 2020 Aug 24. doi:10.7759/cureus.9981

[10] 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. doi:10.1016/j.jpeds.2023.113718

[11] Phua J, Ngerng W, See K, et al. Characteristics and outcomes of culture-negative versus culture-positive severe sepsis. Crit Care. 2013;17(5):R202. Published 2013 Sep 12. doi:10.1186/cc12896

[12] 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(11):4851-4863. doi:10.1128/AAC.00627-10

[13] 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(4):547-553. doi:10.1097/CCM.0000000000002936

[14] Persson I, Östling A, Arlbrandt M, Söderberg J, Becedas D. A Machine Learning Sepsis Prediction Algorithm for Intended Intensive Care Unit Use (NAVOY Sepsis): Proof-of-Concept Study. JMIR Form Res. 2021;5(9):e28000. Published 2021 Sep 30. doi:10.2196/28000

[15] Yao RQ, Jin X, Wang GW, et al. A Machine Learning-Based Prediction of Hospital Mortality in Patients With Postoperative Sepsis. Front Med (Lausanne). 2020;7:445. Published 2020 Aug 11. doi:10.3389/fmed.2020.00445

[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. Published 2023 Apr 3. doi:10.3389/fimmu.2023.1140755

[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(1):635. Published 2023 Sep 27. doi:10.1186/s12879-023-08614-0

[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):1. Published 2023 Jan 3. doi:10.1038/s41597-022-01899-x

[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(3):870. Published 2023 Jan 21. doi:10.3390/jcm12030870

[20] Münch MM, Peeters CFW, Van Der Vaart AW, Van De Wiel MA. Adaptive group-regularized logistic elastic net regression. Biostatistics. 2021;22(4):723-737. doi:10.1093/biostatistics/kxz062

[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(6):410-415. doi:10.1136/emermed-2020-210401

[22] Ince C, Mayeux PR, Nguyen T, et al. THE ENDOTHELIUM IN SEPSIS. Shock. 2016;45(3):259-270. doi:10.1097/SHK.0000000000000473

[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 Jun 1;12(3):603-13. doi: 10.5114/aoms.2015.51208. Epub 2015 Apr 28. PMID: 27279855; PMCID: PMC4889677.

[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(11):e1063-e1143. doi:10.1097/CCM.0000000000005337

[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(2):464-475. Published 2024 Mar 30. doi:10.5114/aoms/172160

[26] Aya HD, Ster IC, Fletcher N, Grounds RM, Rhodes A, Cecconi M. Pharmacodynamic Analysis of a Fluid Challenge. Crit Care Med. 2016;44(5):880-891. doi:10.1097/CCM.0000000000001517

[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(1):110. Published 2023 Mar 13. doi:10.1186/s13054-023-04397-7

[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(9):1138-1147. doi:10.1097/CCM.0000000000005886

[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(9):2357. Published 2023 Sep 20. doi:10.3390/microorganisms11092357

[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(5):1300-1309. doi:10.1213/ANE.0000000000004072

[31] Martin GS, Mannino DM, Moss M. The effect of age on the development and outcome of adult sepsis. Crit Care Med. 2006;34(1):15-21. doi:10.1097/01.ccm.0000194535.82812.ba

[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(1):7-13. doi:10.5603/AIT.2015.0002

[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(2):215-223. doi:10.1097/SHK.0000000000001612

[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(1):143-151.e7. doi:10.1016/j.jtcvs.2019.11.137

[35] Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet. 1974;2(7872):81-84. doi:10.1016/s0140-6736(74)91639-0

[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(1):11-20. doi:10.1056/NEJMoa1611391

[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

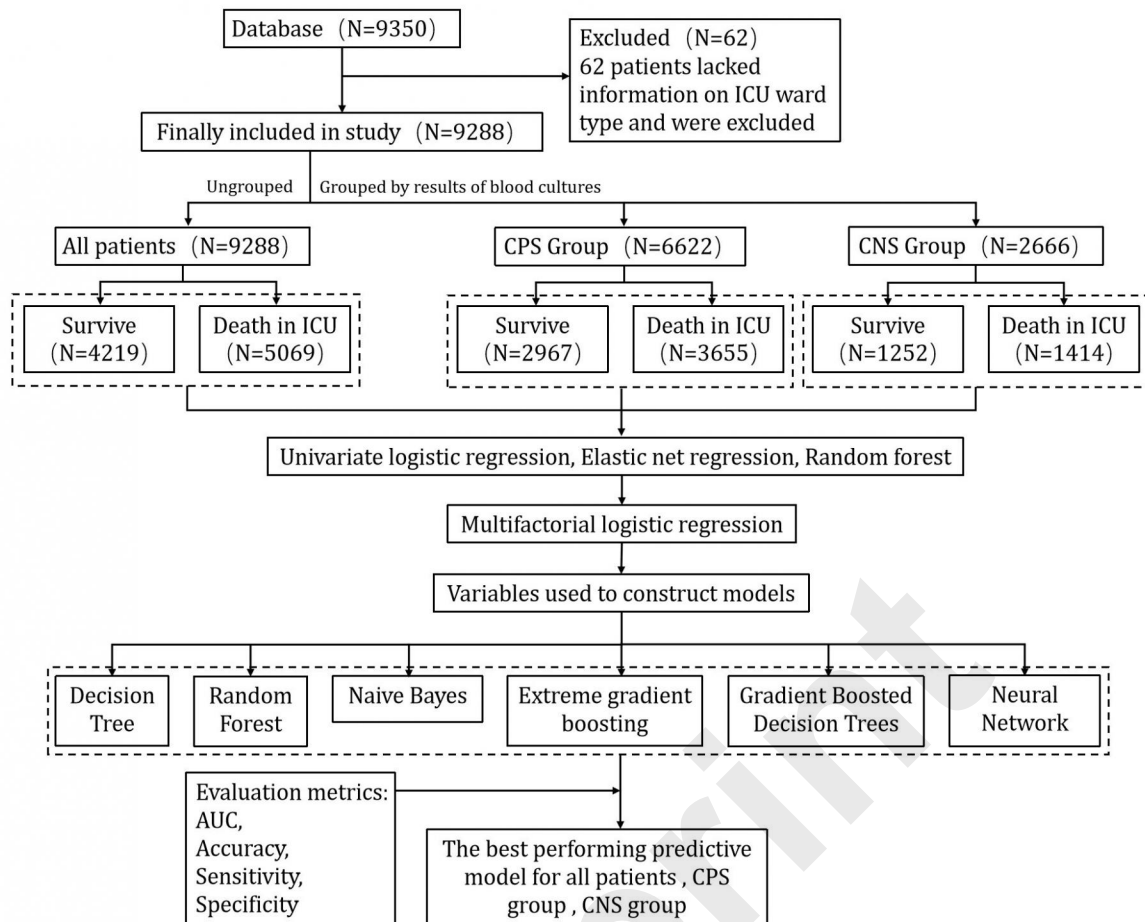Table 1 Comparison of baseline characteristics between the two subgroups in ICU

Variables	CPS group(N=6622)	CNS group(N=2666)	Z/χ^2	P value
age	65.48±16.34	65.52±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	0,7
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)		

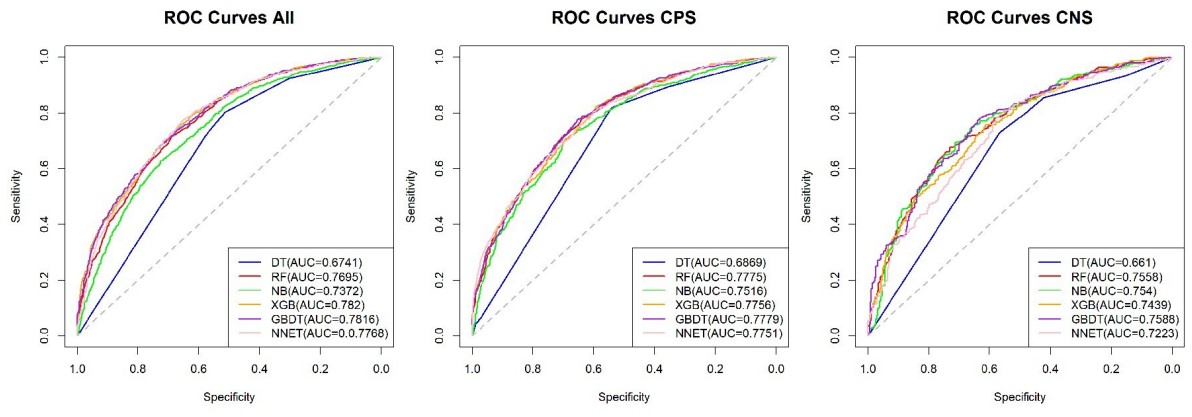| Severity score/(Md[IQR])				
SOFA	7.73 (3.72, 11.75)	7.04 (3.29, 10.8)	82,83	<0.01
SAPSII	42.35 (27.62, 57.07)	40.27 (26.22, 54.31)	80,92	<0.01
CCI	6.17 (3.26, 9.09)	6.06 (3.09, 9.04)	76,79	0,15
Comorbidity/(n,%)				
chronic pulmonary disease	2067 (31.21)	775 (29.07)	4,02	0,05
liver disease	571 (8.62)	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)	0,4	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
90day deaths	2206(33.31)	811(30.42)	7,12	<0.01

Predictive results of six machine learning models based on test sets for septic patients die in

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

```
                    ┌─────────────────────────┐         ┌──────────────────────────────┐
                    │  Database  (N=9350)      │         │ Excluded  (N=62)             │
                    └─────────────────────────┘         │ 62 patients lacked           │
                                 │                       │ information on ICU ward      │
                                 │─────────────────────▶ │ type and were excluded       │
                                 ▼                       └──────────────────────────────┘
                    ┌─────────────────────────┐
                    │ Finally included in study  (N=9288) │
                    └─────────────────────────┘
                 Ungrouped │  Grouped by results of blood cultures
```
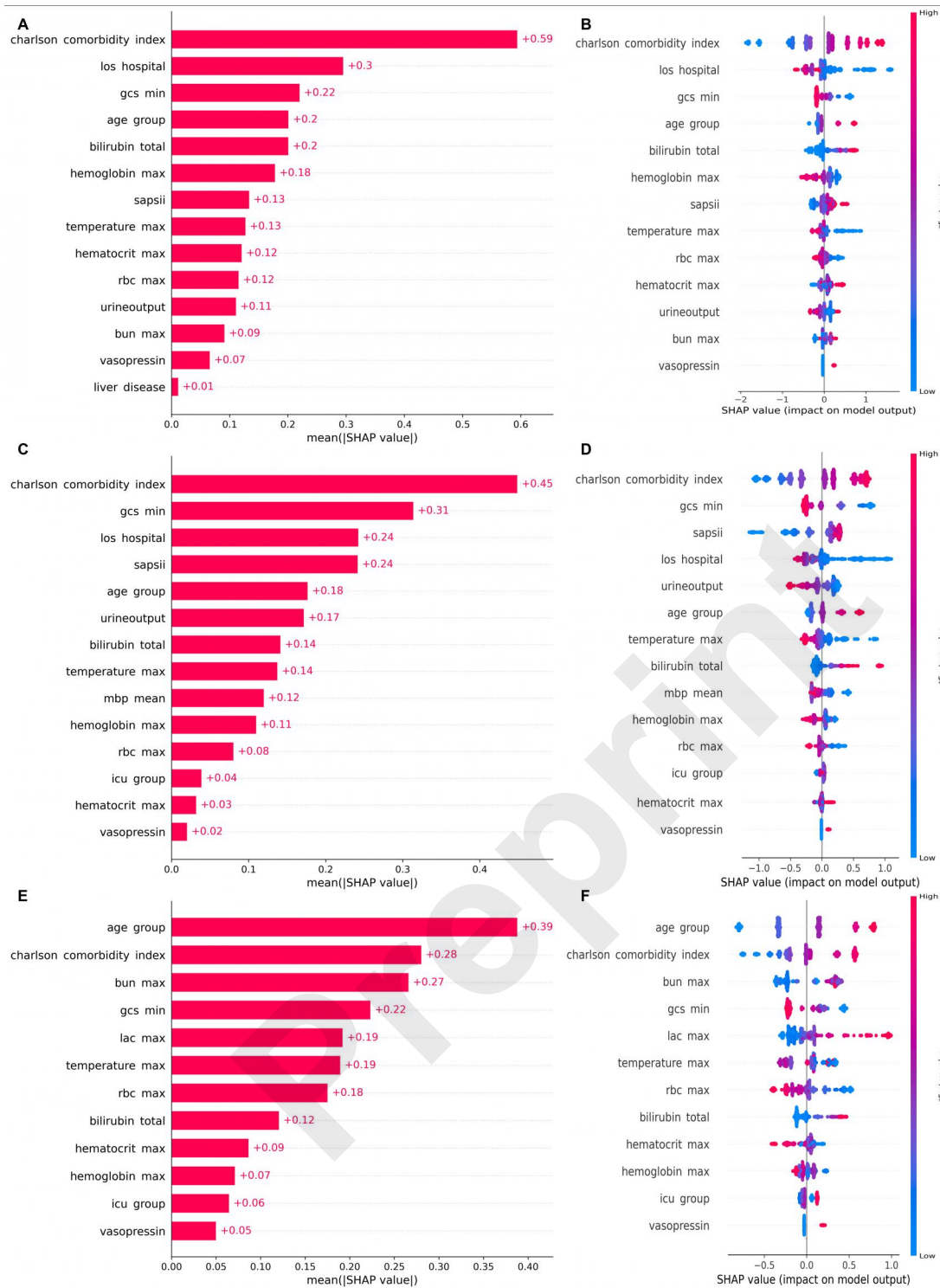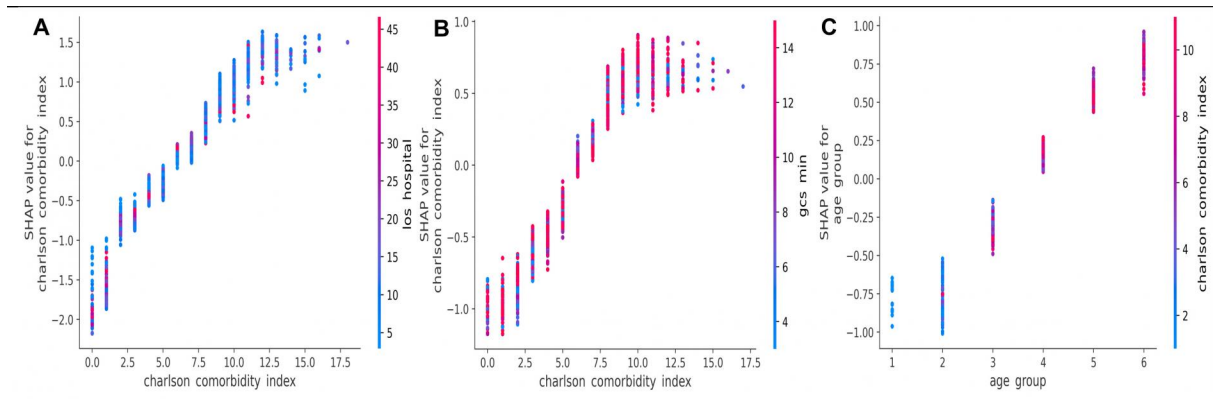

Study design

Summary AUC of 6 machine learning algorithms

Summary of SHAP

Summary of variable interaction dependency plot