Machine learning prediction of early hypothermia in sepsis patients

Ji Li[1], Yifeng Cheng[2]*, Wenxue Tang[2], Yunyan Jing[2], Yali Chen[2], Shixiao Tu[3]

[1]Hangzhou Normal University Affiliated Hospital Zhejiang University School of Medicine Second Affiliated Hospital, China
[2]Hangzhou Normal University Affiliated Hospital, China
[3]The Third People's Hospital of Deqing, China

***Corresponding author:**
Yifeng Cheng
Hangzhou Normal University Affiliated Hospital, China
E-mail: 724661969@qq.com

Submitted: 3 June 2025; **Accepted:** 17 August 2025
Online publication: 29 October 2025

Arch Med Sci
DOI: https://doi.org/10.5114/aoms/209620

Abstract

Introduction: Sepsis is a systemic inflammatory response syndrome caused by infection and remains a leading cause of mortality worldwide. Abnormal body temperature, especially hypothermia (body temperature < 36°C), is a key clinical feature in sepsis patients and is closely associated with disease severity, impaired immune function, and poor prognosis. Early prediction of hypothermia is crucial for timely intervention and improving prognosis.
Material and methods: This study used machine learning algorithms to train and validate a prediction model for early temperature changes in critically ill sepsis patients. Data were extracted from the MIMIC-IV database and five models were established: XGBoost, LR, SVM, KNN, and ANN.
Results: The XGBoost model demonstrated the best predictive performance with AUC values of 0.92 in the training cohort and 0.98 in the validation cohort. \
Conclusions: This model can assist clinicians in identifying sepsis patients at high risk for early hypothermia and implementing early intervention to reduce mortality.

Key words: hypothermia, sepsis, machine learning, prediction model, MIMIC-IV database.

Introduction

Sepsis is a systemic inflammatory response syndrome triggered by infection, often accompanied by organ dysfunction, and is one of the leading causes of high morbidity and mortality worldwide [1–3]. Abnormal body temperature is a key clinical feature of sepsis, including both fever and hypothermia [4]. Compared to fever, hypothermia occurs less frequently in sepsis but carries more significant clinical implications [5]. Studies have shown that hypothermia in sepsis patients is often associated with impaired immune function, metabolic disturbances, and microcirculatory dysfunction [6]. More importantly, hypothermia is closely related to disease severity and poor prognosis, including higher in-hospital mortality and organ dysfunction [7].

Despite the significant harm hypothermia poses to sepsis patients, its early identification and prediction remain a major clinical challenge. Traditional physiological scoring systems (such as SOFA, APACHE II) and laboratory indicators are typically based on static information [8, 9], making

it difficult to comprehensively capture the dynamic process of hypothermia onset, which may result in delayed risk assessment [10]. Furthermore, the mechanisms underlying hypothermia are complex, involving multiple factors such as the source of infection, inflammatory response, immune regulation, patient-specific factors, and therapeutic interventions, which further complicate prediction [11].

With the rapid development of medical big data and artificial intelligence technologies, machine learning (ML) offers a new solution for predicting sepsis-related hypothermia. ML can integrate multimodal patient data (such as physiological parameters, laboratory indicators, and medical history) and, through complex algorithms, uncover hidden patterns to build efficient prediction models [12, 13]. Compared to traditional methods, ML models can dynamically and in real-time capture potential risks, offering higher sensitivity and specificity [14], thereby providing opportunities for early intervention.

The primary objective of this study was to develop a ML-based model that accurately predicts the risk of early hypothermia in sepsis patients. Additionally, this study aimed to identify key factors influencing early hypothermia onset and provide a clinically actionable tool to facilitate timely intervention and improve patient prognosis.

Material and methods

Data source

This study used a large-scale intensive care database for model training, specifically the Medical Information Mart for Intensive Care IV (MIMIC-IV), version 3.0. MIMIC-IV is a database that includes data from all ICU and emergency department patients at Beth Israel Deaconess Medical Center from 2008 to 2019. The database contains patient vital signs, medications, laboratory measurements, provider-recorded observations and notes, fluid balance, procedure codes, diagnostic codes, imaging reports, length of stay, survival data, and more. To access the database, the author (Li Ji) passed the human research participant protection review and obtained certification (ID: 59720689). The MIMIC-IV database was accessed using PostgreSQL software (Version 14.5-1) and Navicat Premium 15, and structured query language (SQL) was used to extract data for the training cohort. The external validation cohort was manually collected from the ICU of three comprehensive hospitals in Zhejiang Province, including sepsis patients diagnosed between January 1, 2023, and December 31, 2023.

Participants

According to the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis 3.0), sepsis is a life- threatening organ dysfunction caused by infection and impaired host response, characterized by an increase of 2 or more points in the Sequential Organ Failure Assessment (SOFA) score. We included 2,623 sepsis patients in the training cohort and 599 sepsis patients in the validation cohort. Patients were excluded if they were under 18 years old, were admitted to the ICU for less than 24 h, or had pre-existing hypothermia (body temperature < 36.0°C) upon admission.

Data extraction and processing

Patient data from the first 24 h after admission were retrieved from the MIMIC-IV database [15] for training, and data from sepsis patients diagnosed in the ICU of Zhejiang University School of Medicine Second Affiliated Hospital, Hangzhou Normal University Affiliated Hospital, and the Third People's Hospital of Deqing between January 1, 2023, and December 31, 2023, were collected for validation. To facilitate the practical implementation and promotion of the prediction model, variables were selected based on their early availability and easy accessibility. A total of 29 input variables were used in this study, categorized as follows:

- Demographic features: sex, age, body mass index (BMI);
- Vital signs: temperature, heart rate, respiratory rate, mean arterial pressure, 24-hour temperature variability at admission;
- Treatment and clinical management: mechanical ventilation within 24 h of admission, use of vasopressors, continuous renal replacement therapy, antibiotic treatment, timing and types of antibiotics, total fluid infusion within 24 h of admission, red blood cell transfusion volume, and plasma transfusion volume;
- Laboratory parameters: white blood cell count, absolute lymphocyte count, absolute neutrophil count, hemoglobin, and blood lactate at admission;
- SOFA score: with only the initial test value included for analysis.

Missing data handling

Variables with missing data are common in the MIMIC-IV, and directly eliminating patients with missing values or analyzing variables with missing values will cause bias. We excluded variables with more than 20% of values missing. For variables with less than 5% of values missing, in the case of continuous variables with normal distribution, the missing values were replaced with the mean for the patient group; in the case of continuous variables with skewed distributions, the missing values were replaced with their median. Multiple imputa-

tion can impute each missing value with multiple plausible possible values. This method takes into account uncertainty behind the missing value and can produce several datasets from which parameters can be estimated, and these coefficients are combined to give an effective estimate of the coefficients. For variables with more than 5% of values missing, we used multiple imputation with the 'mice' package in R to handle the data.

Given the potential for class imbalance in our dataset, we addressed this issue using synthetic oversampling techniques. Specifically, we applied the Synthetic Minority Over-sampling Technique (SMOTE) to the training set. This technique generates synthetic samples for the minority class (patients who developed hypothermia) to balance the number of instances in each class. By doing so, we aimed to prevent the model from being biased towards the majority class and to improve the overall performance of the predictive models.

Statistical analysis

Patients were divided into two groups based on whether they developed hypothermia within 24 h of admission. For continuous variables, values are presented as mean ± standard deviation (SD) if normally distributed, or as median ± interquartile range (IQR) if non-normally distributed. Categorical variables are expressed as total count (percentage). Continuous variables were compared using the *t*-test or Wilcoxon rank-sum test, while proportions were compared using the χ^2 test or Fisher's exact test, as appropriate.

Feature selection and model construction

We utilized the Boruta algorithm, a feature selection method based on random forest theory, to identify truly important features from the given feature set while filtering out those with an insignificant impact. After feature selection, we constructed models using five ML algorithms: Extreme Gradient Boosting (XGBoost), Logistic Regression (LR), k-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Artificial Neural Networks (ANN).

To prevent overfitting, we implemented 10-fold cross-validation during training. This method involves dividing the training data into ten subsets, using nine subsets for training and one subset for validation in each iteration. The process is repeated ten times, with each subset serving as the validation set once. The final model performance is then averaged across all iterations. Additionally, we applied L1 and L2 regularization techniques to Logistic Regression and XGBoost to penalize overly complex models and reduce the risk of overfitting. For XGBoost and ANN, we implemented early stopping to halt training when the validation performance began to degrade.

Hyperparameter tuning was performed using grid search with cross-validation. This method systematically evaluates a range of hyperparameter values to identify the combination that yields the best performance for each model. The specific hyperparameters and their ranges are detailed in the supplementary material.

For model validation, we assessed the performance of the models on an independent validation set using metrics including the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The best-performing model was selected based on its performance on the validation set. Additionally, we designed a simplified nomogram for predicting the occurrence of hypothermia to facilitate clinical application. All analyses were performed using R software.

Results

Baseline characteristics

A total of 27,139 patients were diagnosed with sepsis upon admission, and 24,516 patients were excluded based on the exclusion criteria (Figure 1). Ultimately, 2,623 patients were included in the analysis.

The detailed process of data extraction

The differences in characteristics between sepsis patients who developed hypothermia and those who did not in the training and validation cohorts are shown in Tables I and II. Male patients had a higher likelihood of developing hypothermia within 24 h of ICU admission compared to female patients. Compared to patients who did not develop hypothermia, those who did were older, had lower BMI, temperature, heart rate, and blood pressure at admission, a slower respiratory rate, higher SOFA score, higher blood lactate levels, greater 24-hour temperature variability, higher total fluid infusion, red blood cell transfusion, and platelet transfusion volumes within 24 h. They also had a lower rate of antibiotic use within 24 h, but the antibiotics used were more diverse, with earlier initiation of antibiotics after ICU admission. Additionally, mechanical ventilation, vasopressor, and continuous blood purification therapy were used less frequently ($p < 0.05$).

Feature selection

The feature selection results based on the Boruta algorithm are shown in Figure 2, sorted by *Z*-scores. The 16 variables most closely associated with the occurrence of hypothermia are age,

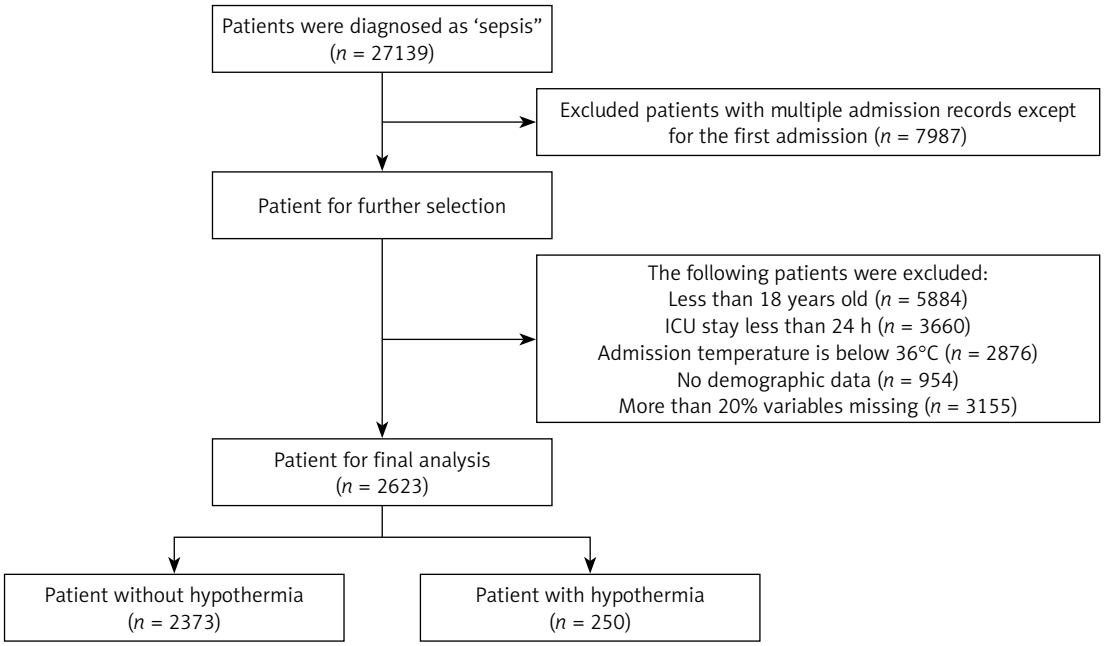

Figure 1. The detailed process of data extraction

Table I. Baseline characteristics of training cohort

Characteristics	Total (*n* = 2623)	Non-hypothermia (*n* = 2373)	Hypothermia (*n* = 250)	*P*-value
Male, *n* (%)	1609 (61.3)	1449 (90.06)	160 (9.94)	0.364
Age [years] mean (SD)	65.14 ±15.37	64.89 ±15.43	67.46 ±14.61	0.012
BMI [kg/m^2] mean (SD)	29.43 ±7.34	29.57 ±7.40	28.07 ±6.53	0.001
First temperature [°C] mean (SD)	36.95 ±0.60	36.98 ±0.59	36.69 ±0.58	0.001
First heart rate [beats/minute] mean (SD)	92.44 ±21.52	93.08 ±21.45	86.38 ±21.30	0.001
First mean BP [mm Hg] mean (SD)	81.34 ±18.15	81.58 ±18.15	79.03 ±18.00	0.034
First respiratory rate [breaths/min] mean (SD)	20.14 ±6.87	20.33 ±6.92	18.37 ±6.17	0.001
SOFA score, mean (SD)	6.95 ±3.90	6.82 ±3.82	8.20 ±4.44	0.001
Total infusion volume in 24 h [ml] mean (SD)	2455.36 ±2239.08	2421.32 ±2208.92	2778.50 ±2489.06	0.016
Total red blood cell transfusion in 24 h [ml], mean (SD)	197.84 ±681.32	170.48 ±622.92	457.52 ±1056.76	0.001
Total plasma transfusion in 24 h [ml] mean (SD)	56.67 ±225.03	47.41 ±195.77	144.52 ±399.48	0.001
WBC [10^9/l] mean (SD)	14.98 ±12.85	15.03 ±13.03	14.50 ±10.99	0.536
Neutrophils, mean (SD)	11.83 ±8.34	11.78 ±8.17	12.26 ±9.82	0.392
Lymphocytes, mean (SD)	1.89 ±10.32	1.89 ±10.71	1.86 ±5.47	0.964
Hemoglobin [g/dl]	10.87 ±2.37	10.44 ±2.35	9.67 ±2.34	0.001
Lactate [mmol/l] mean (SD)	2.38 ±1.84	2.33 ±1.72	2.84 ±2.67	0.003
24-hour temperature variation, mean (SD)	0.014 ±0.070	0.01 ±0.05	0.04 ±0.17	0.022
Time of first antibiotic use, mean (SD)	15.70 ±32.32	16.30 ±33.22	9.95 ±21.28	0.001
Kinds of antibiotic, mean (SD)	1.81 ±1.23	1.80 ±1.24	1.96 ±1.17	0.047
Vasopressor, *n* (%)	1086 (41.4)	955 (87.94)	131 (12.06)	0.001
Continuous renal replacement therapy, *n* (%)	91 (3.5)	55 (60.44%)	36 (39.56%)	0.001
Mechanical ventilation, *n* (%)	1450 (55.3)	1275 (87.93)	175 (12.07)	0.001
Antibiotic treatment, *n* (%)	2203 (84.0)	1973 (89.56)	230 (10.44)	0.001

Table II. Baseline characteristics of validation cohort

Characteristics	Total (n = 599)	Non-hypothermia (n = 473)	Hypothermia (n = 126)	P-value
Male, n (%)	392 (65.44)	361 (92.10)	31 (7.91)	0.923
Age [years] mean (SD)	68.07 ±15.93	68.22 ±15.95	69.51 ±15.90	0.658
BMI [kg/m^2] mean (SD)	22.37 ±4.20	22.56 ±4.08	21.66 ±4.57	0.032
First temperature [°C] mean (SD)	37.20 ±0.88	37.08 ±0.82	37.05 ±0.79	0.001
First heart rate [beats/min] mean (SD)	102.54 ±22.29	103.52 ±21.64	102.29 ±22.48	0.584
First mean BP [mm Hg] mean (SD)	85.72 ±16.38	86.78 ±16.06	81.75 ±17.02	0.001
First respiratory rate [breaths/min] mean (SD)	20.04 ±6.07	20.13 ±6.03	20.04 ±6.09	0.882
SOFA score, mean (SD)	10.49 ±3.89	9.96 ±3.74	12.48 ±3.81	0.001
Total infusion volume in 24 h [ml], mean (SD)	1262.23 ±1089.99	1185.57 ±1002.38	1550.01 ±1335.70	0.001
Total red blood cell transfusion in 24 h [ml], mean (SD)	62.23 ±172.97	29.21 ±120.98	208.73 ±248.52	0.001
Total plasma transfusion in 24 h [ml] mean (SD)	113.19 ±310.48	82.07 ±280.04	230.00 ±348.25	0.001
WBC [10^9/l] mean (SD)	13.17 ±8.86	13.21 ±8.82	13.01 ±9.04	0.819
Neutrophils, mean (SD)	12.01 ±10.52	11.91 ±8.20	12.39 ±16.59	0.749
Lymphocytes, mean (SD)	0.81 ±1.09	1.01 ±1.91	0.75 ±0.73	0.150
Hemoglobin [g/dl]	9.90 ±2.81	10.10 ±2.70	9.13 ±3.09	0.001
Lactate [mmol/l] mean (SD)	3.40 ±6.69	2.65 ±3.50	6.20 ±12.56	0.001
24-hour temperature variation, mean (SD)	0.03 ±0.19	0.02 ±0.04	0.04 ±0.22	0.03
Time of first antibiotic use, mean (SD)	5.37 ±5.93	5.38 ±6.26	5.33 ±4.50	0.931
Kinds of antibiotic, mean (SD)	1.36 ±0.67	1.34 ±0.67	1.44 ±0.65	0.128
Vasopressor, n (%)	359 (59.93%)	264 (73.54%)	95 (26.46%)	0.001
Continuous renal replacement therapy, n (%)	128 (21.37)	49 (38.28)	79 (61.72)	0.001
Mechanical ventilation, n (%)	422 (70.45)	319 (75.59)	103 (24.41)	0.001
Antibiotic treatment, n (%)	594 (99.17)	469 (78.96)	125 (21.04)	0.955

temperature at admission, heart rate, respiratory rate, white blood cell count, absolute neutrophil count, absolute lymphocyte count, hemoglobin, blood lactate, total fluid infusion within 24 h, total red blood cell transfusion within 24 h, total plasma transfusion within 24 h, 24-hour temperature variability, continuous renal replacement therapy (CRRT) treatment, mechanical ventilation, and the timing of antibiotic use.

Feature selection based on the Boruta algorithm

Feature selection based on the Boruta algorithm is shown here. The x-axis represents the names of the variables, and the y-axis represents their Z-scores. The boxplot displays the Z-scores of each variable during the model computation process. The green boxes represent the top 16 important variables, yellow indicates provisional attributes, and red denotes unimportant variables. MV: mechanical ventilation, BMI: body mass Index, Kantibiotic: kinds of antibiotic, MAP: mean arterial pressure, LYM: absolute lymphocyte count, RR: respiratory rate, Tantibiotic: time of first antibiotic use, WBC: white blood cell count,

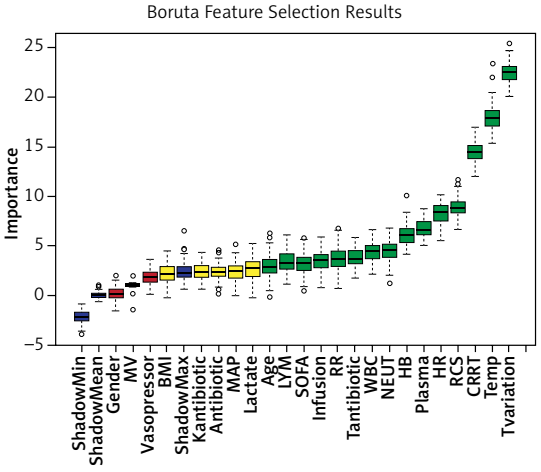

Figure 2. ?????????????????????

NEUT: absolute neutrophil count, HB: hemoglobin, Plasma: total plasma transfusion in 24 h, HR: heart rate, RCS: total red blood cell transfusion in 24 h, CRRT: continuous renal replacement therapy, Temp: temperature, Tvariation: 24-hour temperature variation.

Model performance comparison

We constructed five ML models to predict the occurrence of hypothermia in sepsis patients within 24 h of ICU admission. Figure 3 shows the ROC curve discrimination performance of the five models in both the training and validation cohorts. Among the five models in the training cohort, the XGBoost model (AUC = 0.92) performed best in predicting hypothermia in sepsis patients, followed by SVM (AUC = 0.91), Logistic Regression (AUC = 0.75), KNN (AUC = 0.54), and ANN (AUC = 0.52). Among the five models in the validation cohort, the XGBoost model (AUC = 0.98) again performed best in predicting hypothermia in sepsis patients, followed by SVM (AUC = 0.96), Logistic Regression (AUC = 0.92), KNN (AUC = 0.71), and ANN (AUC = 0.59). Table III presents a set of detailed performance metrics for the five models. In the training cohort, the XGBoost model demonstrated the best discrimination with an AUC of 0.92, accuracy of 0.93, specificity of 0.84, and NPV of 0.98, ranking first. Sensitivity was 0.83, and PPV was 0.34. In the validation cohort, the XGBoost model also demonstrated the best discrimination, with an AUC of 0.98, accuracy of 0.95, specificity of 0.94, and NPV of 0.98, ranking first. Sensitivity was 0.93, and PPV was 0.80.

Nomogram model construction

A nomogram model was constructed based on the variables identified through Boruta analysis

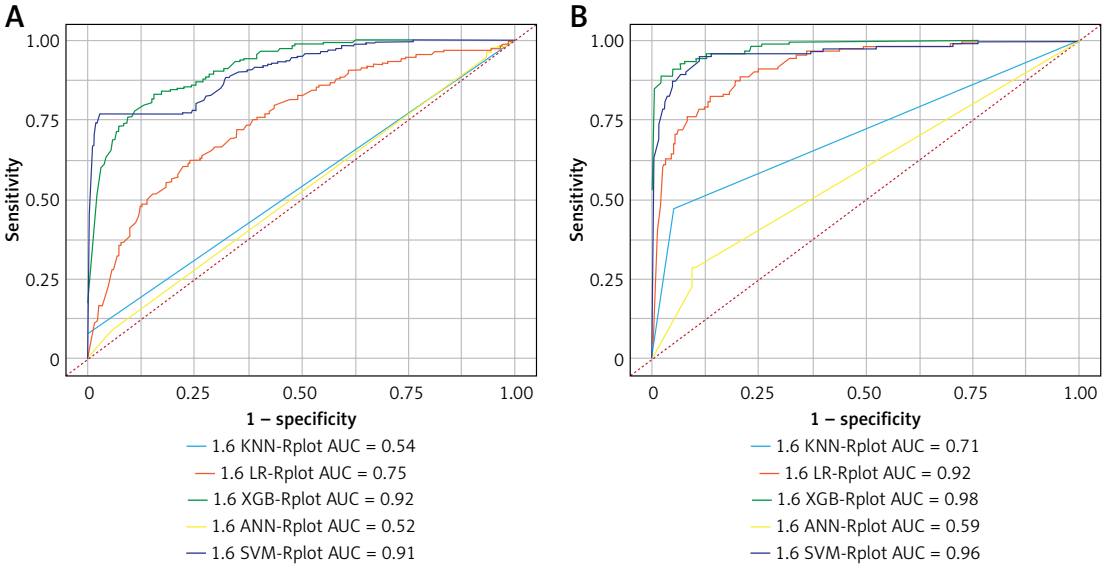

Figure 3. Plots (**A**) and (**B**) show the ROC curves of the five models in the training and validation groups

Table III. Analysis of sensitivity and specificity

Variable	Model	Accuracy	Sensitivity	Specificity	NPV	PPV	AUC
Training sets XGB	XGBoost	0.93	0.83	0.84	0.98	0.34	0.92
	LR	0.92	0.99	0.03	0.43	0.91	0.75
	SVM	0.91	0.99	0.03	0.43	0.91	0.91
	KNN	0.92	0.99	0.08	0.71	0.92	0.54
	ANN	0.91	1	0	0.43	0.91	0.52
Validation sets	XGBoost	0.95	0.93	0.94	0.98	0.80	0.98
	LR	0.90	0.96	0.67	0.81	0.92	0.92
	SVM	0.90	0.96	0.67	0.81	0.79	0.96
	KNN	0.85	0.95	0.46	0.71	0.87	0.71
	ANN	0.8	1	0.06	0.89	0.80	0.59

PPV – positive predictive value, NPV – negative predictive value, AUC – area under the curve, XGBoost – Extreme Gradient Boosting, LR – logistic regression, SVM – Support Vector Machine, KNN – k-Nearest Neighbors, ANN – Artificial Neural Network.

(Figure 4). Each independent variable on the nomogram is assigned a point by drawing a line from the independent variable scale to the point scale (e.g., an age of 20 years is assigned 0 points). The total score is calculated by summing the points assigned to each independent variable on the nomogram. The final score for the occurrence of hypothermia is calculated by drawing a line from the total point scale to the bias risk scale. The score can be used to predict the likelihood of early hypothermia in sepsis patients. The higher the total score is, the greater is the likelihood of hypothermia. Clinicians and nurses can use these readily available indicators to assess the risk of hypothermia in a visual, personalized, and quantitative manner.

Discussion

This study used data from the MIMIC database and applied various ML techniques to construct a risk prediction model for early hypothermia in sepsis patients, which was subsequently validated using ICU data from three domestic centers. Our study suggests that the key variables influencing the occurrence of hypothermia include patient age, temperature at admission, heart rate, respiratory rate, white blood cell count, absolute neutrophil count, absolute lymphocyte count, hemoglobin, blood lactate, 24-hour temperature variability, total fluid infusion within 24 h, total red blood cell transfusion, total plasma transfusion, duration of antibiotic use, and whether continuous blood purification therapy was administered. The ML models effectively predicted the risk of early hypothermia in hospitalized sepsis patients, with the XGBoost model demonstrating the best performance among all models.

Sepsis is a systemic inflammatory response triggered by infection, with body temperature being a key indicator of this response [16, 17]. When hyperthermia fails to effectively alleviate or control sepsis and the condition worsens, it may result in a decrease in body temperature, leading to hypothermia. This often indicates disease deterioration, such as shock or multiple organ failure [18–20]. Temperature variability refers to the degree of fluctuation in body temperature over a specific period. Patients with greater temperature fluctuations often experience disruptions in immune system responses or metabolic instability [6, 11]. Excessive temperature fluctuations

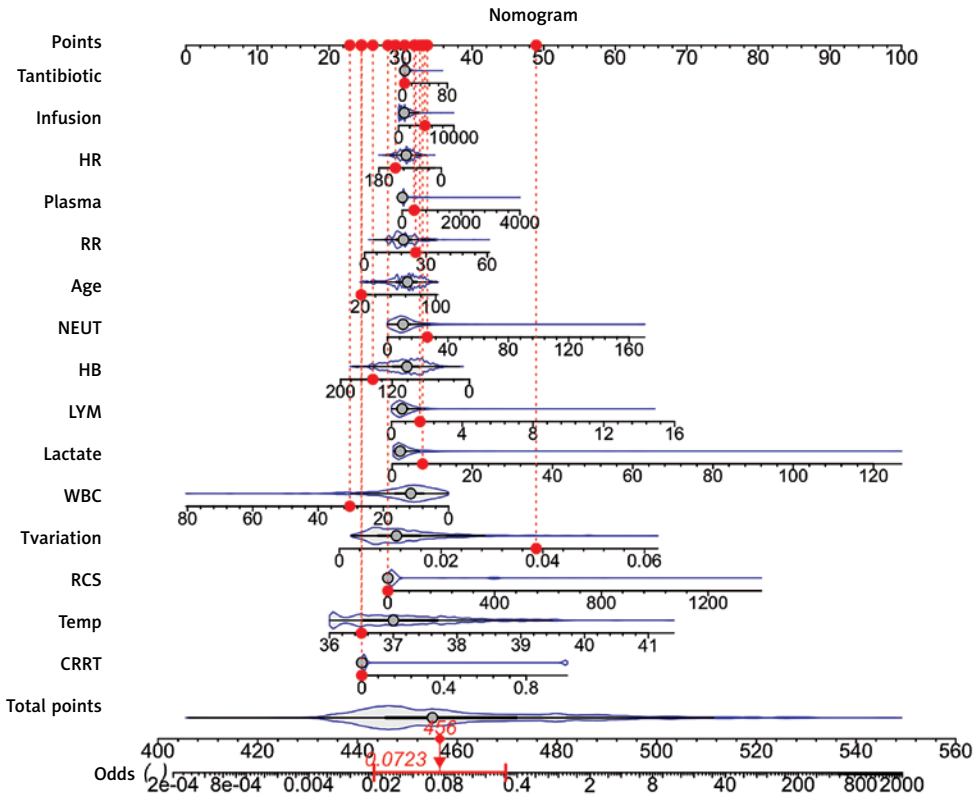

Figure 4. ???????

MV – mechanical ventilation, BMI – body mass index, Kantibiotic – kinds of antibiotic, MAP – mean arterial pressure, LYM – absolute lymphocyte count, RR – respiratory rate, Tantibiotic – time of first antibiotic use, WBC – white blood cell count, NEUT – absolute neutrophil count, HB – hemoglobin, Plasma – total plasma transfusion in 24 hours, HR – heart rate, RCS – total red blood cell transfusion in 24 h, CRRT – continuous renal replacement therapy, Temp – temperature, Tvariation – 24-hour temperature variation.

may lead to rapid energy depletion, making it difficult to maintain normal body temperature levels, thereby increasing the risk of hypothermia [21]. In sepsis patients, the heart rate is typically elevated, reflecting sympathetic nervous system activation and enhanced systemic inflammation [22]. In the later stages of sepsis, an excessively rapid heart rate may lead to heart failure and poor circulation, thereby impairing temperature regulation and increasing the likelihood of hypothermia [23]. An increased heart rate may indicate a reduced compensatory capacity of the body, particularly when accompanied by hemodynamic instability, which increases the risk of hypothermia [24]. Increased respiratory rate is a common manifestation of sepsis, suggesting that the body is attempting to compensate for metabolic acidosis or hypoxia, particularly during shock or multiple organ failure, when blood lactate levels rise [25]. Elevated lactate levels reflect tissue hypoxia, metabolic dysregulation, and cellular dysfunction. High lactate levels cause metabolic disturbances, affecting energy metabolism and heat production, leading to reduced thermogenesis and increasing the risk of hypothermia [26]. In sepsis, white blood cell count is often elevated, reflecting the immune response. An increase in neutrophils is commonly associated with bacterial infections and may trigger a systemic inflammatory response (SIRS) [27]. Excessive inflammatory responses may suppress the function of the temperature-regulating center, leading to a decrease in body temperature rather than an increase, thereby increasing the risk of hypothermia. Moreover, severe leukopenia or functional suppression (e.g., in late-stage sepsis) may weaken the immune response, subsequently affecting normal temperature regulation [28]. Lymphocyte count typically decreases in sepsis patients, especially in immunosuppressed states, where the ability of lymphocytes to combat infection diminishes, further impairing temperature regulation. A reduction in lymphocytes is also closely associated with immune dysregulation in sepsis, potentially leading to temperature response disturbances and the onset of hypothermia [29, 30]. Sepsis patients often present with anemia, and low hemoglobin levels may impair oxygen delivery, particularly under hypoxic conditions, where the body may be unable to maintain body temperature through normal metabolic thermogenesis [31]. Anemia may also exacerbate circulatory dysfunction in sepsis, reducing temperature regulation capacity and increasing the risk of hypothermia [32].

Sepsis patients have a higher risk of hypothermia during bundled treatments such as fluid resuscitation, blood transfusion, antibiotic therapy, and continuous blood purification. Sepsis patients require fluid resuscitation and blood transfusion to restore blood volume, correct anemia, restore hemoglobin levels, and improve oxygen delivery. During fluid resuscitation, the fluids used are typically at room temperature (e.g., saline, crystalloids, or colloids), and in the intensive care unit, the temperature of the fluids may be lower due to environmental conditions. During large-volume and rapid infusion of fluids and blood, the body attempts to adapt by adjusting blood flow and metabolism. If the fluid temperature is too low, it may impair normal temperature regulation [33]. If sepsis patients develop acute kidney injury (AKI), they may require continuous renal replacement therapy (CRRT) to remove metabolic waste and excess fluid. During CRRT, blood passing through the filter typically cools slightly. Although devices often include heating mechanisms, some may cause blood temperature to drop during the purification process, potentially leading to hypothermia. CRRT not only removes waste products but also depletes essential electrolytes, leading to fluctuations in metabolic function, which may indirectly affect the maintenance of body temperature [34, 35].

Age is an important factor influencing temperature regulation in sepsis patients and has a significant impact on the early occurrence of hypothermia. Studies show that the incidence of hypothermia in patients over 65 years old reaches 30–40%, and prognosis is often poor [36]. This may be related to several factors. First, the aging population experiences an immune system decline, with slower initiation of the inflammatory response, resulting in a less noticeable increase in body temperature, enhanced anti-inflammatory responses, and suppression of the fever mechanism, making them more prone to hypothermia [37]. Additionally, the basal metabolic rate gradually decreases with age, and the function of major thermogenic organs, such as the liver and muscles, weakens, impairing heat production ability. At the same time, with aging, the sensitivity of the hypothalamic temperature-regulating center decreases, and the response speed slows. The ability to thermoregulate in response to infection decreases, and the body cannot quickly respond to sepsis-induced stress by increasing metabolic rate. Finally, elderly individuals have poor microcirculation, with reduced ability to redistribute blood flow, making it difficult to maintain organ perfusion and thermal balance [38]. Therefore, as age increases, the probability of early hypothermia increases. This also serves as a reminder that sepsis in elderly patients should be closely monitored, and preventive measures against hypothermia should be implemented [39].

The prediction model for early hypothermia in sepsis patients holds significant clinical potential.

It relies on readily available clinical variables and can be integrated into existing electronic health record (EHR) systems for real-time monitoring and alerting. To ensure its real-world effectiveness, we propose several validation strategies: prospective validation in diverse populations across different regions and healthcare systems; validation in various clinical settings beyond the ICU; long-term follow-up to assess its impact on patient outcomes; and implementation in clinical practice to evaluate its effectiveness in hypothermia prevention and management. Future work should focus on comprehensive validation and seamless integration into clinical workflows to enhance sepsis management and improve patient outcomes.

This study has several limitations. First, we acknowledge that the data in the MIMIC-IV database are incomplete, particularly with the potential absence of key clinical variables (e.g., SCRP, PCT), which may affect the accuracy of the model's predictions. Future research should include more comprehensive data to improve the model and further validate it to enhance its accuracy. Secondly, since this was a retrospective study, the data primarily come from ICU patients, which may introduce selection bias and limit the generalizability of the model in other clinical settings. Therefore, we recommend that future studies use multicenter data to validate this model, to reduce selection bias and increase its generalizability. Additionally, the imbalance between hypothermic and non-hypothermic patients in the dataset may affect the model's performance in predicting mortality.

In conclusion, our study developed a ML model using the XGBoost algorithm to predict early hypothermia in sepsis patients, achieving an AUC of 0.98 in the validation cohort. The model leverages patient demographics, vital signs, laboratory parameters, and clinical treatments as key predictors. While our findings demonstrate the model's strong predictive performance, we acknowledge limitations such as potential selection bias from the MIMIC-IV and three ICUs, class imbalance in the dataset, and the need for further validation in independent datasets. The nomogram provides a useful tool for clinicians, but its real-world effectiveness requires prospective evaluation. Our study underscores the potential of ML for early risk assessment but emphasizes the need for further validation and careful clinical application.

Acknowledgments

The authors would like to thank MIMIC-IV for open access to their database. The opinions expressed in this study are those of the authors and do not represent the opinions of the Beth Israel Deaconess Medical Center.

Data availability

The datasets generated and/or analyzed during the current study are available in the MIMIC database, which is accessible through PhysioNet. To access the data, researchers must complete CITI Program training and sign a Data Use Agreement (DUA). External verification data can be obtained by contacting the corresponding author.

Funding

This work was supported by Huzhou Municipal Science and Technology Bureau (2022GY51, 2023GY76).

Ethical approval

All procedures in this study were conducted in accordance with the ethical standards of the Chinese Human Experimentation Ethics Committee and the 1975 Declaration of Helsinki. This study was reviewed and approved by the Institutional Review Boards of Zhejiang University School of Medicine Second Affiliated Hospital, Hangzhou Normal University Affiliated Hospital, and The Third People's Hospital of Deqing in July 2023, with approval numbers (2023) Lun Shen Yan No. (0769), 2023(E2)-KS- 120, and 2023-EC-040.

Due to the retrospective nature of the study, the Institutional Review Boards of the Zhejiang University School of Medicine Second Affiliated Hospital, Hangzhou Normal University Affiliated Hospital, and Deqing County Third People's Hospital waived the need for obtaining informed consent

Conflict of interest

The authors declare no conflict of interest.

References

1. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016; 315: 801-10.
2. Hotchkiss RS, Moldawer LL, Opal SM, et al. Sepsis and septic shock. Nat Rev Dis Primers 2016; 2: 16045.
3. Machado FR, Cavalcanti AB, Bozza FA, et al. The epidemiology of sepsis in Brazilian intensive care units (the Sepsis PREvalence Assessment Database, SPREAD): an observational study. Lancet Infect Dis 2017; 17: 1180-9.
4. Annane D. Body temperature in sepsis: a hot topic. Lancet Respir Med 2018; 6: 162-3.
5. Kushimoto S, Gando S, Daizoh Saitoh D, et al. The impact of body temperature abnormalities on the disease severity and outcome in patients with severe sepsis: an analysis from a multicenter, prospective survey of severe sepsis. Crit Care 2013; 17: R271.
6. Liu D, Huang SY, Sun JH, et al. Sepsis-induced immunosuppression: mechanisms, diagnosis and current treatment options. Mil Med Res 2022; 9: 56.
7. Ito Y, Kudo D, Kushimoto S. Association between low body temperature on admission and in-hospital mortal-

ity according to body mass index categories of patients with sepsis. Medicine 2022; 101: e31657.
8. Vincent J, de Mendonca A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Crit Care Med 1998; 26: 1793-800.
9. Thakur R, Naga Rohith V, Arora JK. Mean SOFA score in comparison with APACHE II score in predicting mortality in surgical patients with sepsis. Cureus J Med Sci 2023; 15: e36653.
10. Huang Y, JIang S, Li W, et al. Establishment and effectiveness evaluation of a scoring system-RAAS (RDW, AGE, APACHE II, SOFA) for sepsis by a retrospective analysis. J Inflamm Res 2022; 15: 465-74.
11. Sonna LA, Hawkins L, Lissauer ME, et al. Core temperature correlates with expression of selected stress and immunomodulatory genes in febrile patients with sepsis and noninfectious SIRS. Cell Stress Chaperones 2010; 15: 55-66.
12. Handelman GS, Kok HK, Chandra RV, et al. eDoctor: machine learning and the future of medicine. J Intern Med 2018; 284: 603-19.
13. Deo RC. Machine learning in medicine. Circulation 2015; 132: 1920-30.
14. Toh TS, Dondelinger F, Wang D. Looking beyond the hype: applied AI and machine learning in translational medicine. Ebiomedicine 2019; 47: 607-15.
15. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data 2023; 10: 1.
16. Speiser JL, Miller ME, Tooze J, Ip E. A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl 2019; 134: 93-101.
17. Lee BH, Inui D, Suh GY, et al. Association of body temperature and antipyretic treatments with mortality of critically ill patients with and without sepsis: multi-centered prospective observational study. Crit Care 2012; 16: R33.
18. Peres Bota D, Lopes Ferreira F, Lot CM, Vincent JL. Body temperature alterations in the critically ill. Intens Care Med 2004; 30: 811-6.
19. Wiewel MA, Harmon MB, van Vught LA, et al. Risk factors, host response and outcome of hypothermic sepsis. Crit Care 2016; 20: 328.
20. Laupland KB, Zahar JR, Adrie C, et al. Determinants of temperature abnormalities and influence on outcome of critical illness. Crit Care Med 2012; 40: 145-51.
21. Young PJ, Bellomo R. Fever in sepsis: is it cool to be hot? Crit Care 2014; 18: 109.
22. Carrara M, Ferrario M, Bollen Pinto B, Herpain A. The autonomic nervous system in septic shock and its role as a future therapeutic target: a narrative review. Ann Intensive Care 2021; 11: 80.
23. Leibovici L, Gafter-Gvili A, Paul M, et al. Relative tachycardia in patients with sepsis: an independent risk factor for mortality. QJM-Int J Med 2007; 100: 629-34.
24. Ning Y, Li WJ, Lu X, et al. Association between heart rate and mortality in patients with septic shock: an analysis revealed by time series data. BMC Infect Dis 2024; 24: 1088.
25. Broccard A. Prone position, high-frequency oscillatory ventilation, and Hippocrates in acute respiratory distress syndrome. Crit Care Med 2005; 33: 2407-8.
26. Ahmad S, Beckett M. Recovery from pH 6.38: lactic acidosis complicated by hypothermia. Emerg Med J 2002; 19: 169-71.
27. Sibbald WJ, Doig G, Inman KJ. Sepsis, SIRS and infection. Intens Care Med 1995; 21: 299-301.
28. Yaguchi A, Yuzawa J, Klein DJ, Takeda M, Harada T. Combining intermediate levels of the Endotoxin Activity Assay (EAA) with other biomarkers in the assessment of patients with sepsis: results of an observational study. Crit Care 2012; 16: R88.
29. Zhang G, Wang T, An L, et al. U-shaped correlation of lymphocyte count with all-cause hospital mortality in sepsis and septic shock patients: a MIMIC-IV and eICU-CRD database study. Int J Emerg Med 2024; 17: 101.
30. Yang J, Ma B, Tong H. Lymphocyte count trajectories are associated with the prognosis of sepsis patients. Crit Care 2024; 28: 399.
31. Sheng S, Li A, Zhang C, et al. Association between hemoglobin and in-hospital mortality in critically ill patients with sepsis: evidence from two large databases. BMC Infect Dis 2024; 24: 1450.
32. Peng H, Su Y, Luo J, Ding N. Association between admission hemoglobin level and prognosis in sepsis patients based on a critical care database. Sci Rep 2024; 14: 5212.
33. Booke M, Sielenkämper A. Massivtransfusion mit dem Rapid Infusion System. Der Anaesthesist 2001; 50: 926-9.
34. Challener D, Frank R, Kashani K, O'Horo J. Body temperature trends of critically Ill patients on continuous renal replacement therapy: a single-center retrospective study. Am J Med Sci 2022; 364: 404-8.
35. Morsch CMF, Haas JS, Plotnick R, et al. Hypothermia related to continuous renal replacement therapy: incidence and associated factors. Crit Care Sci 2021; 33: 111-8.
36. Warmerdam M, Stolwijk S, Boogert A, et al. Initial disease severity and quality of care of emergency department sepsis patients who are older or younger than 70 years of age. PLoS One 2017; 12: e185214.
37. Michels EHA, Butler J, Reijnders TDY, et al. Association between age and the host response in critically ill patients with sepsis. Crit Care 2022; https://doi.org/10.21203/rs.3.rs-2131537/v1
38. 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.
39. Innocenti F, Tozzi C, Donnini C, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med 2017; 13: 405-12.