Contents lists available at [ScienceDirect](www.sciencedirect.com/science/journal/24058440)

Heliyon

journal homepage: www.cell.com/heliyon

Machine learning based clinical prediction model for 1-year mortality in Sepsis patients with atrial fibrillation

Hong Meng $a,b,c,1}$, Liang Guo $a,b,c,1}$, Yucheng Pan $a,b,c,1}$, Bin Kong a,b , c, Wei Shuai^{a,b,c,**}, He Huang^{a,b,c,*}

^a *Department of Cardiology, Renmin Hospital of Wuhan University, Wuhan, 430060, Hubei, PR China* ^b *Hubei Key Laboratory of Cardiology, Wuhan, 430060, Hubei, PR China*

^c *Cardiovascular Research Institute of Wuhan University, Wuhan, 430060, Hubei, PR China*

ARTICLE INFO

5© CelPress

Keywords: Atrial fibrillation Sepsis Mortality Nomogram MIMIC

ABSTRACT:

Background: Atrial fibrillation (AF) emerges as a pivotal risk determinant for unfavorable outcomes in septic patients. Despite its recognized role, the enduring impact of AF on sepsis prognosis remains ambiguous. This investigation seeks to elucidate the connection between AF and both short and long-term outcomes in sepsis patients. Additionally, it aims to formulate a prognostic model for 1-year mortality utilizing pertinent clinical variables.

Methods: A retrospective analysis encompassed sepsis patients admitted to Beth Israel Deacon Medical Center's intensive care unit. The evaluation encompassed the prevalence of AF and its influence on hospitalization duration, stays in the Intensive Care Unit (ICU), and mortality rates at distinct intervals. Propensity score matching was implemented to mitigate confounding factors. Machine learning techniques, including the Least Absolute Selection and Shrinkage Operator (LASSO) regression and random forest, were deployed for model development.

Results: AF exhibited a correlation with heightened mortality rates at 7 days, 28 days, and 1 year. The resultant predictive model demonstrated superior efficacy compared to prevailing clinical critical illness scores in forecasting mortality risk. Crucial predictors in the model included variables such as RDW, weight, age, BUN, lactate, temperature, MCHC, MBP, ALP, and hemoglobin. *Conclusions:* AF emerges as a substantial peril for adverse outcomes in sepsis patients. The risk model, encompassing pertinent clinical variables, outperformed existing clinical critical illness scores in mortality prediction. This model furnishes valuable insights for risk stratification, augmenting prognostic precision in sepsis patients with concomitant AF.

Abbreviation list

AF Atrial Fibrillation CU and the Second Intensive Care Unit Red Cell Distribution Width Red Cell Distribution Width ALP (State of the Material State o Alkaline Phosphatase

(*continued on next page*)

* Corresponding author. Department of Cardiology, Renmin Hospital of Wuhan University, Wuhan, 430060, Hubei, PR China.

** Corresponding author.

E-mail addresses: sw09120@163.com (W. Shuai), huanghe1977@whu.edu.cn (H. Huang). ¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.heliyon.2024.e38730>

Received 8 February 2024; Received in revised form 27 September 2024; Accepted 28 September 2024

Available online 9 October 2024

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(*continued*)

1. Introduction

Sepsis, as defined by the Third International Consensus Definitions for Sepsis and Septic Shock 3 (Sepsis-3), represents a lifethreatening condition resulting from a dysregulated host response to infection [\[1,2\]](#page-11-0). This condition entails systemic infection and has the potential to impact multiple organs and bodily systems [\[3\]](#page-11-0). The assessment of the medical burden associated with sepsis proves challenging due to factors such as its intricate presentation, diverse clinical outcomes, and varying diagnostic criteria [[4](#page-11-0)]. Notably, sepsis has been linked to an augmented risk of death, hospital readmission, and the emergence of new functional limitations within two years following ICU discharge [[5](#page-11-0)]. A comprehensive systematic review and meta-analysis of 129 studies concerning hospitalized patients revealed a global incidence of 6% for hospital-acquired (HA) sepsis, while intensive care unit-acquired (ICU-A) sepsis was reported at 27%. Mortality rates for HA and ICU-A sepsis were documented at 16% and 35%, respectively [\[6\]](#page-11-0).

Atrial fibrillation (AF) represents a prevalent form of irregular heartbeat affecting millions worldwide, significantly contributing to morbidity and mortality due to its association with hemodynamic instability, heart failure, and embolic events [\[7\]](#page-11-0). Research indicates that AF affects approximately 2-4% of the global adult population, with prevalence escalating with age and societal industrialization [\[8,9\]](#page-11-0). As of 2019, approximately 59.7 million individuals worldwide were affected by AF (including atrial flutter) [[10\]](#page-11-0). However, the prevalence and incidence of AF differ across regions [\[11](#page-11-0)]. Among those over 55 years of age of European descent, one in three individuals is affected, with age-adjusted prevalence of 0.60% for men and 0.37% for women [\[12,13](#page-11-0)]. In China, the largest country in Asia, the latest epidemiology shows that the prevalence of atrial fibrillation in the country is 1.6% [\[14](#page-11-0)]. However, the prevalence and incidence of AF in Asian populations are lower than those in North America, with a relative risk of 0.78 [[15\]](#page-11-0). This indicates the urgent need to reduce the burden of atrial fibrillation.

Existing research has underscored the capacity of sepsis to induce arrhythmias, with AF emerging as the most prevalent among them [\[16,17](#page-11-0)]. Moreover, sepsis is correlated with a six-fold higher risk of AF development, typically manifesting within the initial three days of hospital admission [[18\]](#page-11-0). Sepsis concomitant with AF is characterized by elevated mortality rates and prolonged hospital stays [[19,20\]](#page-11-0). However, the existing body of literature lacks exploration into the long-term prognosis of patients grappling with sepsis and AF. Hence, the principal objective of our study is to scrutinize the association between AF and the prognosis of sepsis patients, with a focus on evaluating and predicting their prognosis through common and readily accessible clinical indicators.

2. Methods

2.1. Data sources

In this retrospective study, data were sourced from the Medical Information Mart for Intensive Care IV (MIMIC-IV) database, encompassing de-identified health-related data from patients admitted to the intensive care unit at Beth Israel Deaconess Medical Center in Boston, MA, USA, spanning the years 2008 to 2019. The database includes diverse health metrics such as laboratory measurements, medications, and vital signs. Access to this data was granted through an online application, contingent on approval. The study focused on version MIMIC-IV 2.0. Ethical approval, entailing exemptions for informed consent and approval for data sharing, was secured from the Institutional Review Board (IRB) of Beth Israel Deaconess Medical Center [[21,22](#page-11-0)].

2.2. Study design

The inclusion criteria encompassed individuals meeting the following conditions: 1) In-room treatment in the intensive care unit, 2) Age exceeding 18 years, 3) Adherence to the definition and diagnostic criteria of sepsis 3.0, jointly articulated by the American Society of Critical Care Medicine (SCCM) and the European Society of Critical Care Medicine (ESICM), and 4) Presence of AF waveforms in ECG monitoring.

Exclusion criteria were defined as follows: 1) Presence of valvular heart disease, 2) Congenital heart disease, 3) History of cardiac pacemaker implantation, 4) Pregnancy status, 5) ICU stay duration less than 24 hours, and 6) For patients with multiple ICU admission

Table1

The baseline characteristics of included patients

CCU: Cardiovascular Intensive Care Unit; ICU: Intensive Care Unit; MICU: Medical Intensive Care Unit; NCU: Neurological Intensive Care Unit; SICU: Surgical Intensive Care Unit; HR: Heart Rate; MBP: Mean Blood Pressure; RR: Resp Rate; M: Mean; SOFA: Sequential Organ Failure Assessment; LODS: Logistic Organ Dysfunction System; OASIS: Oxford Acute Severity of Illness Score; SAPS II: Simplified Acute Physiology Score IISIRS: Systemic Inflammatory Response Syndrome Scores; RRT: Renal Replacement Therapy; MI: Myocardial Infarct; CHF: Congestive Heart Failure; PVD: Peripheral Vascular Disease; CVD: Cerebrovascular Disease; COPD: Chronic Obstructive Pulmonary Disease

records, only data from the initial admission record were included.

Data extraction involved the retrieval of: 1) Demographic details, including age, gender, and weight; 2) Mean values of vital signs during the 24 hours post-ICU admission, encompassing pulse rate, respiration rate, temperature, blood pressure, and oxygen saturation; 3) Outcomes, such as hospital and ICU stay durations, hospital mortality, 7-day mortality, 28-day mortality, and 1-year mortality; 4) Severity and relative scores, including Sequential Organ Failure Assessment (SOFA), Logistic Organ Dysfunction System (LODS), Simplified Acute Physiology Score II (SAPS II), Oxford Acute Severity of Illness Score (OASIS), systemic inflammatory response syndrome scores (SIRS), and CHA2DS2-VASc scores; 5) Mean or poorest value of laboratory tests during the 24 hours post-ICU admission.

2.3. Statical analysis

The normal distribution of variables was verified using Agostino's test. Presentation of continuous parametric variables involved mean (standard deviation), while non-parametric variables were presented as median (interquartile range). Comparisons for parametric and non-parametric variables utilized Unpaired Student's t-test or Mann–Whitney U-test, respectively. Categorical variables were compared using the χ^2 test or Fisher's exact test.

In order to address the disparities in baseline covariates between the AF group and the non-AF group, we employed propensity score matching (PSM) methodology. Initially, a logistic regression model was constructed with AF status as the dependent variable and various baseline covariates including age, sex, comorbidities, vital signs, and treatment at admission as independent variables. This model was utilized to estimate each patient's propensity score for developing AF - representing the likelihood of a patient being assigned to the AF group based on their baseline characteristics. Subsequently, we utilized the Greedy Nearest Neighbor Algorithm for matching purposes while imposing a 0.05 standard deviation constraint on propensity scores to ensure robustness of matches. A 1:1 matching ratio was then chosen and match quality evaluation was conducted post-matching by comparing differences in baseline covariates between groups.

Multivariate Logistic regression analyzed independent risk factors for sepsis patient prognosis (in-hospital death, 7-day death, 28 day death, 90-day death, and 1-year death). Kaplan–Meier method generated 1-year survival curves, with the log-rank test comparing differences between patients with AF and without AF.

For subsequent prediction model construction, only sepsis patients with AF were included in the follow-up analysis. The data from these patients were randomly divided into a training set (70%) and a validation set (30%). The training set facilitated variable screening and model construction, while the validation set gauged model generalizability and averted overfitting. Least Absolute Selection and Shrinkage Operator (LASSO) regression was utilized for variable selection, followed by a random forest algorithm and cross-validation method for the identification of the most significant variables. Model testing employed a calibration curve of 10,000 bootstrap samples to assess predictive performance. Receiver operating characteristic (ROC) analysis and decision curve analysis (DCA) compared the risk model's performance with commonly used acute and critical scores.

Statistical analyses were conducted using SPSS 27.0 and R 4.2.1. Variables with missing values exceeding 30% were excluded. For remaining missing variables, multiple imputation entailed randomly selecting the mean value from the last five complete cases, generating five imputed datasets. The mean value of these datasets was then used as the missing value for continuous variables, while the most common value among the datasets determined the missing value for categorical variables. A significance level of p *<* 0.05 denoted statistical significance.

3. Results

3.1. Baseline characteristics of included patients

The initial analysis encompassed 7,909 sepsis patients. Median patient age was 69 years (range: 57-80), with 44% being female. Patients with AF were notably older (median age 77 vs. 65 years, p *<* 0.001) and had a higher male proportion (59% vs. 54%, p *<* 0.001) compared to non-AF patients. AF patients exhibited slightly higher body weight (median weight 78 vs. 77 Lb, $p = 0.015$). Vital signs, renal replacement therapy, mechanical ventilation, and medication use (epinephrine, norepinephrine, phenylephrine,

vasopressin) were more frequent in sepsis patients with AF. Comorbidities like congestive heart failure, peripheral vascular disease, cerebrovascular disease, chronic pulmonary disease, diabetes, and renal disease were more prevalent in AF patients. No significant differences were observed in hypertension and liver disease prevalence. Detailed baseline characteristics are presented in [Table 1.](#page-2-0)

3.2. Atrial fibrillation as a risk factor for poor prognosis in Sepsis patients

Patients with AF exhibited significantly elevated rates of in-hospital mortality, 7-day mortality, 28-day mortality, and 1-year mortality. AF patients also experienced prolonged hospital and ICU stays [\(Table 2\)](#page-3-0).

Multivariate logistic regression, considering covariates such as age, sex, weight, marital status, insurance type, scores (SOFA, LODS, SAPS II) on the first ICU day, vital signs, and AF presence, confirmed AF as an independent risk factor for in-hospital mortality, 7-day mortality, 28-day mortality, 90-day mortality, and 1-year mortality (Table 3). In addition, we investigated the effect of atrial fibrillation duration on patient outcomes. Interestingly, AF duration did not affect in-hospital mortality, 7-day mortality, 28-day mortality, 90-day mortality or 1-year mortality in the initial univariate regression analysis (P *>*0.05). This result suggests that, once the sepsis patients with atrial fibrillation, must have poor prognosis.

To eliminate confounding factors as much as possible, we then performed PSM. According to [Table 1](#page-2-0), we found that the baseline characteristics of the patients in the AF group and the patients in the non-AF group were balanced after PSM. In addition, we also visualized the matching process ([Figure 1](#page-5-0)), and the visualization results contained three dimensions. First, the kernel density plot [\(Figure 1A](#page-5-0)) shows the results of the global PS scores of the two groups of patients. Second, a scatter plot, [Figure 1B](#page-5-0), shows the location of each patient's distribution according to the PS score before and after matching. Finally, the love plot [\(Figure 1C](#page-5-0)) shows the patients in the two groups before and after matching for each confounding factor. Two groups of 4082 patients were ultimately matched for further analysis. We found that the overall condition, individual condition, and baseline condition of the two groups of patients were basically similar after matching. The result of PSM also revealed persistently poorer prognosis in AF patients compared to those without AF. Kaplan-Meier analysis consistently showed significantly lower 1-year survival rates in AF patients (p *<* 0.05, [Figure 2](#page-6-0)).

3.3. Development of a risk model for 1-year mortality in Sepsis patients with atrial fibrillation

Given AF's independent risk status, a predictive model based on first-day laboratory tests of sepsis patients with AF was developed. Analyses highlighted specific poor laboratory indices within 24 hours of admission for AF patients, including lactic acid, pH, PO2, PCO2, total CO2, MCHC, MCV, RDW, BUN, calcium, chlorine, creatinine, sodium, potassium, INR, PT, and PTT ([Table 4\)](#page-7-0).

Subsequently, a training set (70%) and a validation set (30%) were randomly derived from 2,475 sepsis patients with AF. LASSO regression and random forest algorithm identified the most significant variables for the predictive model, revealing RDW, weight, age, BUN, lactate, temperature, MCHC, MBP, ALP, and hemoglobin ([Fig. 3, Fig. 4,](#page-8-0) Supplementary Table 1).

The risk model, compared to established scores (SOFA, LODS, SAPS II, OASIS, SIRS, CHA2DS2-VASc), exhibited superior discrimination in both the training set (AUC 0.757) and the validation set (AUC 0.776) for identifying sepsis patients at 1-year mortality risk. Decision curve analysis confirmed the model's superior performance, and calibration curves demonstrated good agreement between predicted and observed probabilities [\(Table 5,](#page-9-0) [Fig. 5](#page-10-0), Supplementary Figure 1, Supplementary Figure 2)

3.4. Visualizing risk prediction models

To enhance interpretability, a nomogram and scoring system were developed. This user-friendly tool allows clinicians to calculate an individual's total score, predicting their 1-year mortality probability. An online calculator based on the nomogram was also provided for real-time variable result viewing ([Fig. 6](#page-10-0), online calculator link: ([https://whurmhmh.shinyapps.io/sepsis/\)](https://whurmhmh.shinyapps.io/sepsis/)).

4. Discussion

Table 3

Our study highlights the significant impact of AF on the short- and long-term outcomes of sepsis patients, including prolonged hospitalization, increased ICU stays, and elevated mortality rates at various time points. Furthermore, we have developed a predictive model that outperforms widely used clinical critical illness scores, offering improved prognostic capabilities.

In our study, we observed a relatively high prevalence of AF among sepsis patients in the ICU (31.29%), which exceeds the rates

Multifactor logistic regression results of atrial fibrillation in different outcomes

β: Regression Coefficients; SE: Standard Error; OR: Odds Ratio; CI: Confidence Interval

Fig. 1. Propensity score matching. (A). The kernel density plot indicates that the propensity scores of the two groups were best matched after adjustment. (B). Point plots indicate the matches for each patient between the two groups as well as those that were not matched. (C). The love point shows the situation of each variable before and after matching, after matching each variable has basically reached the equilibrium state.

reported in previous studies [\[23](#page-11-0),[24](#page-11-0)]. Remarkably, patients with sepsis combined with AF exhibited an in-hospital mortality rate of approximately 46%, consistent with prior research [[15\]](#page-11-0). Moreover, our study investigated the incidence of 7-day death, 28-day death, and 1-year death in sepsis patients, contributing novel insights not explored in previous studies. To ensure reliable comparisons, we employed propensity score matching to eliminate confounding factors that could impact 1-year mortality, resulting in a well-balanced study cohort [[25\]](#page-11-0). By this method, we significantly eliminated the unbalanced baseline characteristics between patients in the AF group as well as the non-AF group and eliminated the confounding factors that may affect 1-year mortality. The Kaplan–Meier's survival analysis was also performed both before and after matching, and the results showed that 1-year mortality was significantly higher in patients with AF than in patients without AF.

During constructing our predictive model, we discovered that the maximum value of lactate played a crucial role as a predictive variable. Elevated lactate levels have been recognized as an essential marker for prognosis in sepsis patients, indicating high metabolic activity and impaired organ function $[26]$ $[26]$. Numerous studies have established the positive correlation between sustained increases in serum lactate and organ dysfunction, as well as the predictive value of lactate for short-term mortality in sepsis patients [27–[30\]](#page-11-0). Although no direct studies have confirmed the association between serum lactate and the risk of death in AF patients, it is worth noting that sepsis-induced myocardial damage and excessive inflammatory response may contribute to the development of AF [\[31,32](#page-12-0)].

In our study, RDW emerged as a significant predictor within the constructed risk model. RDW, a measure of red cell size variability, has been associated with increased mortality in sepsis patients and demonstrated independent prognostic value in patients with AF [\[33](#page-12-0),[34\]](#page-12-0). Basic studies have also shown that RDW leads to severe cellular dysfunction and even multiple organ dysfunction syndrome

Fig. 2. The curves of Kaplan–Meier's survival analysis. (A). The curves of Kaplan–Meier's survival analysis in the unmatched cohort. (B). The curves of Kaplan–Meier's survival analysis in the matched cohort.

(MODS) in sepsis patients through inflammatory as well as oxidative stress pathways [[35,36](#page-12-0)]. Mean corpuscular volume, a measure of red blood cell size, has also been shown to be associated with adverse cardiovascular outcomes, and this relationship may be mediated through RDW [\[37](#page-12-0)].

Blood urea nitrogen (BUN) is a marker of renal function and protein catabolism that can be affected by sepsis. A systematic review and meta-analysis of 16 studies with 10,282 patients found that higher BUN levels were associated with increased mortality in sepsis patients, with a pooled odds ratio of 1.92 (95% confidence interval: 1.64–2.25) [[38\]](#page-12-0). A prospective observational study of 198 patients with sepsis in the ICU found that there was a nonlinear correlation between BUN and 30-day mortality, with a cutoff level of 41.1 mg/dL [[39\]](#page-12-0). A retrospective cohort study 3 of 1,161 patients with sepsis in the ICU found that the BAR was an independent predictor of 28-day mortality, with an adjusted odds ratio of 1.12 (95% confidence interval: 1.07–1.18) per unit increase [\[40](#page-12-0)].

ALP, an enzyme responsible for phosphate group removal, serves as a biomarker for liver and bone diseases. A population-based study using NHANES data demonstrated a relationship between elevated ALP levels and mortality in the general population [[41\]](#page-12-0). Furthermore, research has highlighted the role of ALP in promoting cardiovascular disease and mortality by facilitating vascular calcification, potentially achieved through pyrophosphate hydrolysis in arterial mediators [\[42,43](#page-12-0)]. Inflammation may offer another explanation for the association between elevated serum ALP levels and adverse cardiovascular outcomes and mortality [\[44](#page-12-0)]. A previous study indicated that increased serum ALP levels contribute to the risk of death by influencing serum C-reactive protein, an inflammation marker [[45\]](#page-12-0).

As shown in this study, hemoglobin and MCHC are both important indicators to measure anemia and important variables in the prediction model, and lower values will represent a worse prognosis of patients. Previous studies have confirmed that anemia is a common symptom in patients with sepsis and is highly associated with poor prognosis in patients with AF [\[46,47](#page-12-0)]. Anemia itself may lead to tissue hypoxia, which can exacerbate organ dysfunction and increase the risk of death in septic patients [\[48](#page-12-0)]. Moreover, another study concluded that cell-free hemoglobin and its prosthetic group heme can contribute to organ dysfunction and death. The pathological mechanisms include nitric oxide consumption, vasoconstriction, oxidative injury to lipid membranes, activation of the transcription factor NF-κB, endothelial injury as well as iron-driven oxidative inhibition of glucose metabolism [\[49](#page-12-0)].

Other variables in the model, such as body weight, represent the presence of malnutrition, which may highly affect the prognosis of patients after sepsis. Age itself can affect the survival status of general patients, so it is understandable that age is related to the poor prognosis of patients with sepsis and AF [[24\]](#page-11-0). Low MBP and temperature indicate that patients have respiratory and circulatory system disorders, which may be the potential cause of poor long-term prognosis of patients [\[50](#page-12-0)–52]. Our study also found that thromboplastin time was associated with a high risk of death in patients with sepsis. Thromboplastin time is a routine test used to assess the function of the coagulation system. Several studies have shown that prolonged thromboplastin time is associated with increased mortality in patients with sepsis [[53\]](#page-12-0).

Lastly, we compared our risk model with commonly used clinical scoring systems and demonstrated its superior predictive performance. While existing scoring systems have proven effective in assessing short-term prognosis in sepsis patients through retrospective studies, they often fall short in predicting long-term outcomes [\[54,55](#page-12-0)]. Our risk model, incorporating a comprehensive set of predictive variables, offers improved prognostic accuracy across both short and long-term periods.

Table 4

The characteristics of included patients when first ICU admission

WBC: White blood cell; MCH: Mean corpuscular hemoglobin; MCHC: Mean corpuscular hemoglobin concentration; MCV: Mean corpuscular volume; RBC: Red blood cell; RDW: Red blood cell distribution width; BUN: Blood urea nitrogen; INR: International normalized ratio; PT: Prothrombin time; PTT: Partial thromboplastin time; ALT: Alanine transaminase; ALP: Alkaline phosphatase; AST: Aspartate aminotransferase

5. Clinical implications

This study presents several key strengths that distinguish it from previous research. Firstly, it utilizes an innovative machine learning-based model to predict 1-year mortality in sepsis patients with atrial fibrillation (AF), allowing for more accurate and timely identification of high-risk patients compared to traditional statistical methods. Secondly, the analysis is based on a large, comprehensive dataset from multiple medical centers, enhancing the generalizability and robustness of the findings. The extensive data enabled rigorous propensity score matching, effectively minimizing potential confounding factors. Additionally, by focusing on sepsis patients with AF, the study addresses a critical clinical issue, as AF is a common complication in sepsis and understanding its impact on long-term mortality is crucial for improving patient outcomes. The detailed description of the propensity score matching process and the inclusion of various baseline covariates ensure the reliability and reproducibility of the findings, adding methodological rigor to the study. Moreover, the findings have significant implications for clinical practice, as the ability to predict long-term mortality can guide clinicians in prioritizing interventions for high-risk patients, potentially improving survival rates and optimizing resource allocation in healthcare settings. Finally, the study builds a simple machine learning model and visualizes patient risk in order to identify high-risk patients and enhance their management and treatment, which can inform future research and lead to better management strategies for patients with these conditions. These strengths underscore the significance of this research and its potential impact on patient care and future studies.

6. Study Limitations

To maintain a comprehensive understanding of our findings, it is important to acknowledge the limitations of our study. Firstly, being a single-center retrospective study, our findings may be subject to limitations related to generalizability, particularly considering the predominantly white race population at the Beth Israel Deacon Medical Center in the United States. Secondly, different infection sites can impact the prognosis of sepsis patients, and our study did not differentiate between different infection sites. Lastly, the inability to analyze the dynamics of individual scoring systems limits our understanding of their direct impact on sepsis patient

Fig. 3. Lasso regression analysis process. In the lasso regression process, a larger λ would punish the linear model with more variables more. (A). The variable is continuously removed from the model as λ increases. (B). After LASSO regression screening, two models were obtained. One model is lambda.min, which is the one in which the average of the least objective parameter is obtained out of all the λ values. The other model is lambda.1se, which refers to the λ value of the simplest model within a variance range of lambda.min. (C). The importance ranking of specific variables in lambda.min.

prognosis.

7. Conclusions

To summarize, our study highlights the significance of AF as a prominent risk factor for 1-year mortality among sepsis patients in the ICU. By analyzing relevant variables obtained within the initial 24 hours of ICU admission, such as RDW, weight, age, BUN, lactate, temperature, MCHC, MBP, ALP and hemoglobin, we have developed an accurate predictive model for assessing the 1-year mortality risk in sepsis patients with coexisting AF.

CRediT authorship contribution statement

Hong Meng: Writing – original draft, Conceptualization. **Liang Guo:** Resources, Investigation. **Yucheng Pan:** Software, Investigation. **Bin Kong:** Visualization, Methodology. **Wei Shuai:** Writing – review & editing, Supervision. **He Huang:** Writing – review & editing, Project administration.

Fig. 4. Random forest algorithm results of the model. (A). In the parameter adjustment process of random forest, the error is minimized when the number of decision trees in the forest is 768. (B) and (C). 5-fold cross-validation results of the random forest algorithm and the final result. (D). The variable importance of random forest algorithm finally.

SOFA: Sequential Organ Failure Assessment; LODS: Logistic Organ Dysfunction System; SAPS II: Simplified Acute Physiology Score II; OASIS: Oxford Acute Severity of Illness Score; SIRS: Systemic Inflammatory Response Syndrome Scores; AUC: The area under the curve; CI: Confidence Interval

Fig. 5. The curves of receiver operating characteristic. (A). The curves of receiver operating characteristic in the training set. (B). The curves of receiver operating characteristic in the validation set.

Ethics approval and consent to participate

Ethical approval for data collection and research resource creation was obtained from the Institutional Review Board (IRB) of Beth Israel Deaconess Medical Center, which granted exemptions for informed consent and approved data sharing

Consent for publication

All authors have participated in the work and have reviewed and agree with the content of the article. None of the article contents are under consideration for publication in any other journal or have been published in any journal.

Data availability statement

The authors do not have permission to share data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research presented in this manuscript was supported by grants from the National Natural Science Foundation of China (No. 82070330), which facilitated the execution of this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at [https://doi.org/10.1016/j.heliyon.2024.e38730.](https://doi.org/10.1016/j.heliyon.2024.e38730)

References

- [1] M. Singer, C.S. Deutschman, C.W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, et al., The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3), JAMA 315 (2016) 801–810, [https://doi.org/10.1001/jama.2016.0287.](https://doi.org/10.1001/jama.2016.0287)
- [2] [R. Gauer, D. Forbes, N. Boyer, Sepsis: Diagnosis and Management, Am Fam Physician 101 \(2020\) 409](http://refhub.elsevier.com/S2405-8440(24)14761-5/sref2)–418.
- [3] A.M. Taeb, M.H. Hooper, P.E. Marik, Sepsis: Current Definition, Pathophysiology, Diagnosis, and Management, Nutr Clin Pract 32 (2017) 296-308, [https://doi.](https://doi.org/10.1177/0884533617695243) [org/10.1177/0884533617695243.](https://doi.org/10.1177/0884533617695243)
- [4] W. Wang, C.-F. Liu, Sepsis heterogeneity, World J Pediatr 19 (2023) 919–927,<https://doi.org/10.1007/s12519-023-00689-8>.
- [5] L. Biason, C. Teixeira, J.S. Haas, C. da R. Cabral, G. Friedman, Effects of Sepsis on Morbidity and Mortality in Critically Ill Patients 2 Years After Intensive Care Unit Discharge, Am J Crit Care 28 (2019) 424–432,<https://doi.org/10.4037/ajcc2019638>.
- [6] R. Markwart, H. Saito, T. Harder, S. Tomczyk, A. Cassini, C. Fleischmann-Struzek, et al., Epidemiology and burden of sepsis acquired in hospitals and intensive care units: a systematic review and meta-analysis, Intensive Care Med 46 (2020) 1536–1551, <https://doi.org/10.1007/s00134-020-06106-2>. [7] [C. Gutierrez, D.G. Blanchard, Diagnosis and Treatment of Atrial Fibrillation, Afp. 94 \(2016\) 442](http://refhub.elsevier.com/S2405-8440(24)14761-5/sref7)–452.
- [8] G. Hindricks, T. Potpara, N. Dagres, E. Arbelo, J.J. Bax, C. Blomström-Lundqvist, et al., 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC, Eur Heart J 42 (2021) 373–498, <https://doi.org/10.1093/eurheartj/ehaa612>.
- [9] E.J. Benjamin, P. Muntner, A. Alonso, M.S. Bittencourt, C.W. Callaway, A.P. Carson, et al., Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association, Circulation 139 (2019) e56–e528, [https://doi.org/10.1161/CIR.0000000000000659.](https://doi.org/10.1161/CIR.0000000000000659)
- [10] G.A. Roth, G.A. Mensah, C.O. Johnson, G. Addolorato, E. Ammirati, L.M. Baddour, et al., Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study, J Am Coll Cardiol 76 (2020) 2982–3021, <https://doi.org/10.1016/j.jacc.2020.11.010>.
- [11] H. Li, X. Song, Y. Liang, X. Bai, W.-S. Liu-Huo, C. Tang, et al., Global, regional, and national burden of disease study of atrial fibrillation/flutter, 1990-2019: results from a global burden of disease study, 2019, BMC Public Health 22 (2022) 2015, [https://doi.org/10.1186/s12889-022-14403-2.](https://doi.org/10.1186/s12889-022-14403-2)
- [12] C. Magnussen, T.J. Niiranen, F.M. Ojeda, F. Gianfagna, S. Blankenberg, I. Njølstad, et al., Sex Differences and Similarities in Atrial Fibrillation Epidemiology, Risk Factors, and Mortality in Community Cohorts: Results From the BiomarCaRE Consortium (Biomarker for Cardiovascular Risk Assessment in Europe), Circulation 136 (2017) 1588–1597, [https://doi.org/10.1161/CIRCULATIONAHA.117.028981.](https://doi.org/10.1161/CIRCULATIONAHA.117.028981)
- [13] L. Staerk, B. Wang, S.R. Preis, M.G. Larson, S.A. Lubitz, P.T. Ellinor, et al., Lifetime risk of atrial fibrillation according to optimal, borderline, or elevated levels of risk factors: cohort study based on longitudinal data from the Framingham Heart Study, BMJ 361 (2018) k1453, <https://doi.org/10.1136/bmj.k1453>.
- [14] S. Shi, Y. Tang, Q. Zhao, H. Yan, B. Yu, Q. Zheng, et al., Prevalence and risk of atrial fibrillation in China: A national cross-sectional epidemiological study, Lancet Reg Health West Pac 23 (2022) 100439, <https://doi.org/10.1016/j.lanwpc.2022.100439>.
- [15] G.L. Wells, P.E. Morris, Incidence and Prognosis of Atrial Fibrillation in Patients With Sepsis, Cardiol Res 2 (2011) 293–297,<https://doi.org/10.4021/cr108w>. [16] A. Ayaz, M. Ibrahim, A. Arshad, Long-term Risk of Stroke After New-Onset Atrial Fibrillation in Sepsis Survivors: A 2-Year Follow-Up Study, J Intensive Care
- Med (2024) 8850666241251755, <https://doi.org/10.1177/08850666241251755>. [17] A. Vyas, R. Desai, A. Vasavada, N. Ghadge, A. Jain, D. Pandya, et al., Intersection of sepsis, atrial fibrillation, and severe obesity: a population-based analysis in
- the United States, Int J Obes (Lond). 48 (2024) 224–230,<https://doi.org/10.1038/s41366-023-01403-4>. [18] I. Induruwa, E. Hennebry, J. Hennebry, M. Thakur, E.A. Warburton, K. Khadjooi, Sepsis-driven atrial fibrillation and ischaemic stroke. Is there enough evidence
- to recommend anticoagulation? Eur J Intern Med 98 (2022) 32–36, <https://doi.org/10.1016/j.ejim.2021.10.022>. [19] R. Desai, B. Hanna, S. Singh, A. Omar, A. Deshmukh, G. Kumar, et al., Trends and Outcomes in Sepsis Hospitalizations With and Without Atrial Fibrillation: A
- Nationwide Inpatient Analysis, Crit Care Med 47 (2019) e630-e638, https://doi.org/10.1097/CCM.000000000003806. [20] E.Y. Wang, O.L. Hulme, S. Khurshid, L.-C. Weng, S.H. Choi, A.J. Walkey, et al., Initial Precipitants and Recurrence of Atrial Fibrillation, Circ Arrhythm
- Electrophysiol 13 (2020) e007716, <https://doi.org/10.1161/CIRCEP.119.007716>. [21] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, L.A. Celi, R. Mark, MIMIC-IV, version 2.0, 2022. <https://doi.org/10.13026/7vcr-e114>.
- [22] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, et al., PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (2000) E215–220, <https://doi.org/10.1161/01.cir.101.23.e215>.
- [23] A.J. Walkey, M.A. Greiner, S.R. Heckbert, P.N. Jensen, J.P. Piccini, M.F. Sinner, et al., Atrial fibrillation among Medicare beneficiaries hospitalized with sepsis: incidence and risk factors, Am Heart J 165 (2013) 949–955.e3, [https://doi.org/10.1016/j.ahj.2013.03.020.](https://doi.org/10.1016/j.ahj.2013.03.020)
- [24] R. Desai, B. Hanna, S. Singh, A. Omar, A. Deshmukh, G. Kumar, et al., Trends and Outcomes in Sepsis Hospitalizations With and Without Atrial Fibrillation: A Nationwide Inpatient Analysis, Crit Care Med 47 (2019) e630–e638, [https://doi.org/10.1097/CCM.0000000000003806.](https://doi.org/10.1097/CCM.0000000000003806)
- [25] P.C. Austin, An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies, Multivariate Behav Res 46 (2011) 399–424, <https://doi.org/10.1080/00273171.2011.568786>.
- [26] M. Garcia-Alvarez, P. Marik, R. Bellomo, Sepsis-associated hyperlactatemia, Crit Care 18 (2014) 503, [https://doi.org/10.1186/s13054-014-0503-3.](https://doi.org/10.1186/s13054-014-0503-3)
- [27] L. He, D. Yang, Q. Ding, Y. Su, N. Ding, Association Between Lactate and 28-Day Mortality in Elderly Patients with Sepsis: Results from MIMIC-IV Database, Infect Dis Ther 12 (2023) 459–472, [https://doi.org/10.1007/s40121-022-00736-3.](https://doi.org/10.1007/s40121-022-00736-3)
- [28] M. Singer, C.S. Deutschman, C.W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, et al., The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3), JAMA 315 (2016) 801–810, [https://doi.org/10.1001/jama.2016.0287.](https://doi.org/10.1001/jama.2016.0287)
- [29] Z. Liu, Z. Meng, Y. Li, J. Zhao, S. Wu, S. Gou, et al., Prognostic accuracy of the serum lactate level, the SOFA score and the qSOFA score for mortality among adults with Sepsis, Scand J Trauma Resusc Emerg Med 27 (2019) 51, <https://doi.org/10.1186/s13049-019-0609-3>.
- [30] S.M. Ryoo, J. Lee, Y.-S. Lee, J.H. Lee, K.S. Lim, J.W. Huh, et al., Lactate Level Versus Lactate Clearance for Predicting Mortality in Patients With Septic Shock Defined by Sepsis-3, Crit Care Med 46 (2018) e489–e495, <https://doi.org/10.1097/CCM.0000000000003030>.
- [31] J. Yu, R. Zheng, P. Yang, D. Wang, Construction of a predictive model and prognosis of left ventricular systolic dysfunction in patients with sepsis based on the diagnosis using left ventricular global longitudinal strain, J Intensive Care 10 (2022) 29, [https://doi.org/10.1186/s40560-022-00621-8.](https://doi.org/10.1186/s40560-022-00621-8)
- [32] Y. Kakihana, T. Ito, M. Nakahara, K. Yamaguchi, T. Yasuda, Sepsis-induced myocardial dysfunction: pathophysiology and management, J Intensive Care 4 (2016) 22, <https://doi.org/10.1186/s40560-016-0148-1>.
- [33] C.H. Kim, J.T. Park, E.J. Kim, J.H. Han, J.S. Han, J.Y. Choi, et al., An increase in red blood cell distribution width from baseline predicts mortality in patients with severe sepsis or septic shock, Crit Care 17 (2013) R282, [https://doi.org/10.1186/cc13145.](https://doi.org/10.1186/cc13145)
- [34] W. Saliba, O. Barnett-Griness, G. Rennert, Red cell distribution width and all-cause mortality in patients with atrial fibrillation: A cohort study, J Arrhythm 33 (2017) 56–62, [https://doi.org/10.1016/j.joa.2016.06.001.](https://doi.org/10.1016/j.joa.2016.06.001)
- [35] J.K. Kolls, Oxidative stress in sepsis: a redox redux, J Clin Invest 116 (2006) 860–863, [https://doi.org/10.1172/JCI28111.](https://doi.org/10.1172/JCI28111)
- [36] M. Straat, R. van Bruggen, D. de Korte, N.P. Juffermans, Red blood cell clearance in inflammation, Transfus Med Hemother 39 (2012) 353–361, [https://doi.org/](https://doi.org/10.1159/000342229) [10.1159/000342229](https://doi.org/10.1159/000342229).
- [37] L. Dai, S.L. Mick, K.R. McCrae, P.L. Houghtaling, E.H. Blackstone, C.G. Koch, Prognostic Value of Preoperative Red Cell Distribution Width: Fine-Tuning by Mean Corpuscular Volume, Ann Thorac Surg 108 (2019) 1830–1838,<https://doi.org/10.1016/j.athoracsur.2019.04.072>.
- [38] J. Min, J. Lu, L. Zhong, M. Yuan, Y. Xu, The correlation study between blood urea nitrogen to serum albumin ratio and prognosis of patients with sepsis during hospitalization, BMC Anesthesiology 22 (2022) 404, <https://doi.org/10.1186/s12871-022-01947-4>.
- [39] Y. Wang, S. Gao, L. Hong, T. Hou, H. Liu, M. Li, et al., Prognostic impact of blood urea nitrogen to albumin ratio on patients with sepsis: a retrospective cohort study, Sci Rep 13 (2023) 10013, <https://doi.org/10.1038/s41598-023-37127-8>.
- [40] X. Li, R. Zheng, T. Zhang, Z. Zeng, H. Li, J. Liu, Association between blood urea nitrogen and 30-day mortality in patients with sepsis: a retrospective analysis, Ann Palliat Med 10 (2021) 11653–11663, [https://doi.org/10.21037/apm-21-2937.](https://doi.org/10.21037/apm-21-2937)
- [41] M. Tonelli, G. Curhan, M. Pfeffer, F. Sacks, R. Thadhani, M.L. Melamed, et al., Relation between alkaline phosphatase, serum phosphate, and all-cause or cardiovascular mortality, Circulation 120 (2009) 1784–1792, [https://doi.org/10.1161/CIRCULATIONAHA.109.851873.](https://doi.org/10.1161/CIRCULATIONAHA.109.851873)
- [42] M. Abramowitz, P. Muntner, M. Coco, W. Southern, I. Lotwin, T.H. Hostetter, et al., Serum alkaline phosphatase and phosphate and risk of mortality and hospitalization, Clin J Am Soc Nephrol 5 (2010) 1064–1071,<https://doi.org/10.2215/CJN.08621209>.
- [43] S.G. Wannamethee, N. Sattar, O. Papcosta, L. Lennon, P.H. Whincup, Alkaline phosphatase, serum phosphate, and incident cardiovascular disease and total mortality in older men, Arterioscler Thromb Vasc Biol 33 (2013) 1070–1076, [https://doi.org/10.1161/ATVBAHA.112.300826.](https://doi.org/10.1161/ATVBAHA.112.300826)
- [44] K.A. Lomashvili, P. Garg, S. Narisawa, J.L. Millan, W.C. O'Neill, Upregulation of alkaline phosphatase and pyrophosphate hydrolysis: potential mechanism for uremic vascular calcification, Kidney Int 73 (2008) 1024–1030, [https://doi.org/10.1038/ki.2008.26.](https://doi.org/10.1038/ki.2008.26)
- [45] A. Kerner, O. Avizohar, R. Sella, P. Bartha, O. Zinder, W. Markiewicz, et al., Association between elevated liver enzymes and C-reactive protein: possible hepatic contribution to systemic inflammation in the metabolic syndrome, Arterioscler Thromb Vasc Biol 25 (2005) 193–197, [https://doi.org/10.1161/01.](https://doi.org/10.1161/01.ATV.0000148324.63685.6a) [ATV.0000148324.63685.6a.](https://doi.org/10.1161/01.ATV.0000148324.63685.6a)
- [46] P.C. Hébert, G. Wells, M.A. Blajchman, J. Marshall, C. Martin, G. Pagliarello, et al., A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. Transfusion Requirements in Critical Care Investigators, Canadian Critical Care Trials Group, N Engl J Med 340 (1999) 409–417, <https://doi.org/10.1056/NEJM199902113400601>.
- [47] Z. Zhang, C. Jiang, L. He, Y. Bai, J. Wu, R. Hu, et al., Associations of anemia with death and major bleeding in patients with atrial fibrillation: A report from the Chinese Atrial Fibrillation Registry Study, Clin Cardiol 45 (2022) 91–100, <https://doi.org/10.1002/clc.23764>.
- [48] K. Effenberger-Neidnicht, M. Hartmann, Mechanisms of Hemolysis During Sepsis, Inflammation 41 (2018) 1569–1581, [https://doi.org/10.1007/s10753-018-](https://doi.org/10.1007/s10753-018-0810-y) [0810-y](https://doi.org/10.1007/s10753-018-0810-y).
- [49] Y. Jiang, F.-Q. Jiang, F. Kong, M.-M. An, B.-B. Jin, D. Cao, et al., Inflammatory anemia-associated parameters are related to 28-day mortality in patients with sepsis admitted to the ICU: a preliminary observational study, Annals of Intensive Care 9 (2019) 67, https://doi.org/10.1186/s13613-019-0542-
- [50] M.P. Casaer, D. Mesotten, G. Hermans, P.J. Wouters, M. Schetz, G. Meyfroidt, et al., Early versus late parenteral nutrition in critically ill adults, N Engl J Med 365 (2011) 506–517, [https://doi.org/10.1056/NEJMoa1102662.](https://doi.org/10.1056/NEJMoa1102662)
- [51] R.M. Bateman, M.D. Sharpe, J.E. Jagger, C.G. Ellis, J. Solé-Violán, M. López-Rodríguez, et al., 36th International Symposium on Intensive Care and Emergency Medicine : Brussels, Belgium. 15-18 March 2016, Crit Care 20 (2016) 94,<https://doi.org/10.1186/s13054-016-1208-6>.
- [52] G.D. Brandon, S. Adeniyi-Jones, S. Kirkby, D. Webb, J.F. Culhane, J.S. Greenspan, Are outcomes and care processes for preterm neonates influenced by health insurance status? Pediatrics 124 (2009) 122–127, <https://doi.org/10.1542/peds.2008-1318>.
- [53] L. Chen, X.D. Han, Y.L. Li, C.X. Zhang, X.Q. Xing, [Incidence and risk factors for cardiovascular events in patients hospitalized with community-acquired pneumonia], Zhonghua Xin Xue Guan Bing Za Zhi 48 (2020) 228–235,<https://doi.org/10.3760/cma.j.cn112148-20190617-00342>.
- [54] V. Moreno-Torres, A. Royuela, E. Múñez, A. Ortega, Á. Gutierrez, P. Mills, et al., Better prognostic ability of NEWS2, SOFA and SAPS-II in septic patients, Med Clin (Barc). 159 (2022) 224–229, <https://doi.org/10.1016/j.medcli.2021.10.021>.
- [55] L. Wang, Z. Zhang, T. Hu, Effectiveness of LODS, OASIS, and SAPS II to predict in-hospital mortality for intensive care patients with ST elevation myocardial infarction, Sci Rep 11 (2021) 23887, <https://doi.org/10.1038/s41598-021-03397-3>.