A time series based predictive model for mortality risk among sepsis patients in the intensive care unit

Hao Yang[1#], Jiaxi Li[2#], Chi Zhang[3], Xueli Zhang[4,5], Bairong Shen [3*]

1. Information Center, West China Hospital, Sichuan University, Chengdu, China
2. Department of Clinical Laboratory Medicine, Jinniu Maternity and Child Health Hospital of Chengdu, Chengdu, China
3. Joint Laboratory of Artifcial Intelligence for Critical Care Medicine, Department of Critical Care Medicine and Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, China
4. Guangdong Eye Institute, Department of Ophthalmology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China.
5. Medical Research Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China.

# These authors contributed equally

*Correspondence may also be addressed to Bairong Shen: Phone/Fax: +86-28-85164199; Email: bairong.shen@scu.edu.cn

Funding: This study was supported by the National Natural Science Foundation of China (No. 32070671, 32270690, China)

Abstract: Sepsis constitutes a perilous medical state demanding the precise and prompt identification of patients at elevated risk to optimize therapeutic interventions and prognostic projections. In this study, we propose a time-series modeling approach based on Transformers to predict the outcomes of patients in the ICU. Through the establishment of daily time-series models

based on the patients' duration of ICU stay and the integrating features from preceding days, the model achieves a heightened congruence with the patients' authentic circumstances, thus enabling the intricate delineation of evolving alterations in their health dynamics. Our findings elucidate a direct correlation between the augmentation of integrated data volume and the commensurate enhancement in the model's performance. On average, our model achieved an initial AUC of 0.87 (±0.021) on the first day, which further increases to 0.92 (±0.009) on the fifth day. Additionally, through the utilization of the SHAP visualization algorithm, we reveal activation heatmaps of life features associated with mortality as the time sequence progresses. These results underscore the potential of our approach in providing valuable insights and reliable predictions for patient prognosis and treatment in ICU settings.

Key words: Sepsis; Transformer; Time-series; Visualization; Predicting mortality.

Introduction

Sepsis, an exacerbated systemic reaction in response to infection, possesses the potential to severe medical consequences and even death without timely identification and therapeutic intervention.[1]. While sepsis can impact individuals of all age cohorts, certain populations, including infants, individuals afflicted with chronic ailments, immunocompromised individuals, and the elderly, manifest heightened susceptibility. [1]. In the United States alone, sepsis was responsible for 201,092 deaths in 2019, with approximately three-quarters of these fatalities occurring in individuals aged 65 and above [2].

Accurately predicting mortality among sepsis patients is crucial for clinical physicians, as it facilitates not only the evaluation of disease severity but also the optimization of treatment strategies, reduction adverse outcomes, and extension of patient longevity. Presently, various clinical scoring systems, such as the Sequential Organ Failure Assessment (SOFA) score [3] and the Acute Physiology and Chronic Health Evaluation (APACHE-II) scoring system [4], assist clinicians in evaluating sepsis severity and predicting adverse events. Nonetheless, these scoring systems are devised for the broader population of critically ill patients and do not specifically target sepsis. Given the critical nature of sepsis, the diagnosis and treatment are great important , and an early pre-DIC diagnosis could potentially enhance patient survival rates.

Recently, machine learning in medical research has facilitated the development of predictive

models customized to specific clinical requirements and data characteristics, exhibiting superior predictive performance in anticipating adverse outcomes compared to conventional clinical scoring systems. The construction of existing sepsis mortality prediction models predominantly relies on the utilization of machine learning models., like logistic regression, random forest, and XGBoost, and achieved a good area under the receiver operating characteristic curve (AUC) in the range of 0.70-0.85[5,6,7], there is still sufficient potential to improve their performance.

However, the majority of research predominantly focused on constructing predictive models using data from isolated instances or patient-specific time intervals, inadvertently overlooking the temporal evolution of underlying disease conditions In this study, we exploited the eICU Collaborative Research Database [8], focusing on patients diagnosed with sepsis, and conducted experiments by extracting relevant data such as vital signs and laboratory test results related to sepsis. Trough organizing this data chronologically and analyzing the statistical characteristics and temporal patterns of the time series samples, we constructed a time series model based on the Transformer architecture [9] to predict the mortality risk in sepsis patients. Furthermore, visual algorithms were used to analyze activation characteristics at different time intervals to explore their correlation with the risk of sepsis death. Our objective is to provide novel insights and approaches for the treatment and prognosis of sepsis by establishing a temporal model.

Samples and method

The eICU Collaborative Research Database is a large, multi-center intensive care unit (ICU) database established through a collaboration between the Massachusetts Institute of Technology (MIT) and Philips Group. The database contains clinical data of over 200,000 patients from 208 hospitals in the United States, recorded between 2014 and 2015. It includes demographic information, vital signs, laboratory test results, treatments, diagnoses, and other data. The data quality is high, validated through multiple research studies [8].

The data inclusion criteria: The clinical data of patients with a diagnosis of sepsis in the diagnostic table of the eICU database were collected. Exclusion criteria included patients under 18 years of age, patients with unknown gender, the diagnosis was invalid upon discharge, patients with an ICU stay duration of less than 24 hours, and patients with more than 30% missing data.

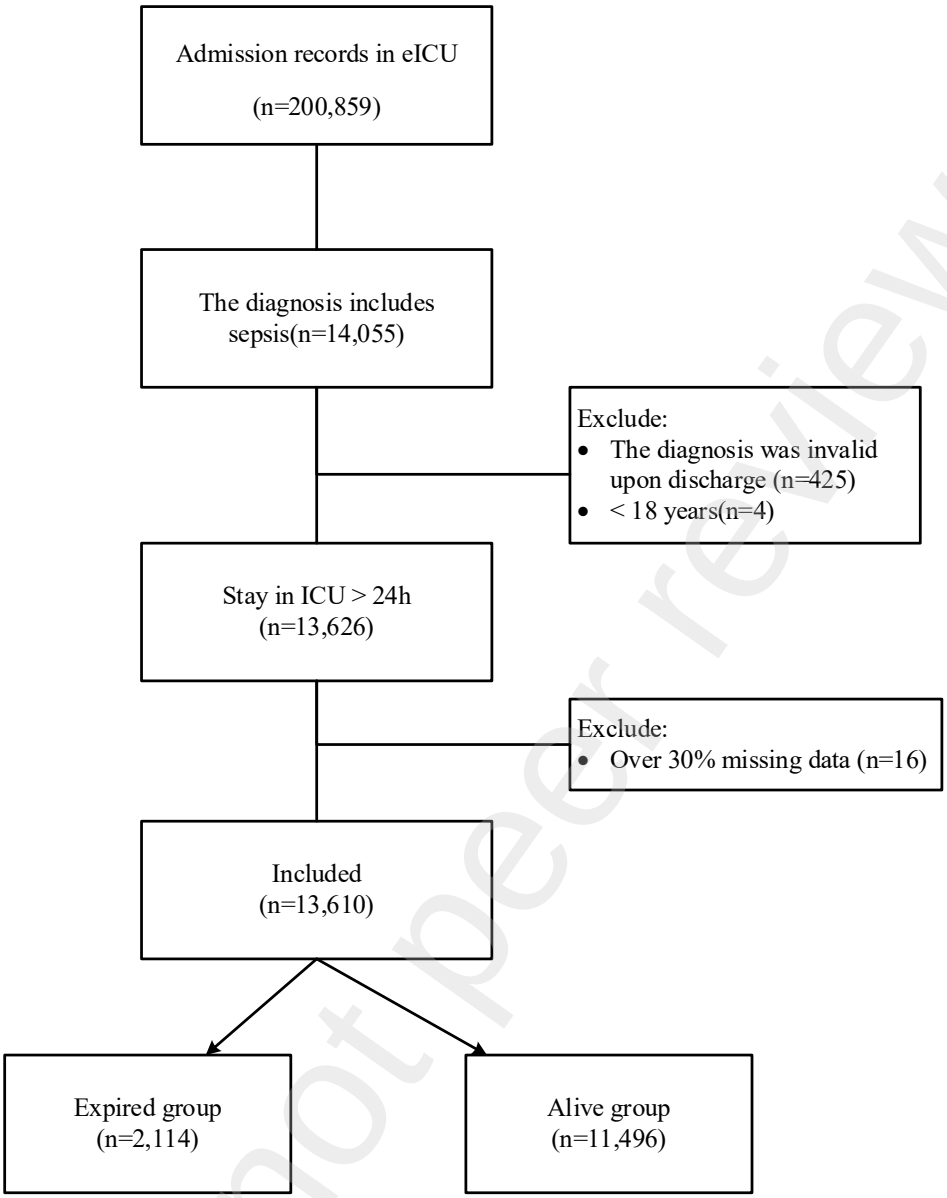

Figure1. Data inclusion and exclusion process

Data collection

Drawing upon clinical expertise, published literature, and data available in the eICU database, we gathered the following categories of data: 1) demographic information encompassing gender, age, and ethnicity; 2) vital signs recorded post-ICU admission, including heart rate, mean arterial pressure, respiratory rate, and oxygen saturation; 3) laboratory test results obtained after ICU admission, such as creatinine and hemoglobin levels, among others. In total, there are 226 distinct vital sign data features (refer to Appendix 1 for further details).

Data Preprocessing

Using Python data preprocessing libraries such as NumPy and pandas [10, 11], the data will be

organized chronologically based on the admission timeline to the ICU, utilizing the "offset" field in each eICU data table. With 24 data sampling points per day, representing hourly intervals, a total of 226 features will be considered. For each hourly sampling point, the nearest available data value will be selected, resulting in a time-series matrix of size 24 x× 226. Subsequently, predictive models will be constructed for different ICU stay durations, such as Day 1, Day 2... Day 5, depending on the length of stay at admission. The data will be randomly divided into training, validation, and testing sets in a ratio of 7:2:1 to ensure robust model performance.

Missing data and filling

In time series analysis, missing data poses a common challenge. Recognizing the limitations of traditional missing data imputation methods and the presence of missing data in our study, we opted to employ the forward imputation method and random forest imputation [12]. These methods were chosen to address the specific needs of our time series analysis and to effectively handle missing data.

Baseline

A total of 13,610 patients were selected from the eICU database for inclusion in the study. Among them, 2,114 patients died during their ICU stay, while 11,496 patients survived. The data extraction was performed using the PostgreSQL database system [13]. Statistical analysis was conducted using SPSS 22.0 [14], while data cleaning, model construction, and performance evaluation were carried out using Python. Continuous variables were expressed as median (interquartile range), while categorical data were presented as counts (percentages). Mann-Whitney U test [15] was used for analyzing continuous variables, and chi-square test [16] was employed to examine significant differences in categorical variables.

Table 1. Baseline characteristics of ICU sepsis patients

Characteristic		ALIVE	EXPIRED	p-test
ICU Stays (mean (SD))		4.171 (5.340)	5.518 (7.933)	<0.001
apachescore (mean (SD))		66.020 (23.845)	88.479 (30.147)	<0.001
gender (%)	Female	5693 (49.52)	1004 (47.49)	0.1889
	Male	5801 (50.46)	1110 (52.51)	
	Unknown	2 (0.02)	0 (0.00)	

age (mean (SD))		65.252 (16.309)	69.807 (13.987)	<0.001
ethnicity (%)		78 (0.68)	15 (0.71)	0.4997
	African American	1136 (9.88)	222 (10.50)	
	Asian	175 (1.52)	43 (2.03)	
	Caucasian	9080 (78.98)	1641 (77.63)	
	Hispanic	425 (3.70)	88 (4.16)	
	Native American	94 (0.82)	17 (0.80)	
	Other/Unknown	508 (4.42)	88 (4.16)	
height (mean (SD))		168.341 (14.618)	168.067 (14.465)	0.428
Unit type (%)	Cardiac ICU	741 (6.45)	155 (7.33)	0.0143
	CCU-CTICU	697 (6.06)	121 (5.72)	
	CSICU	132 (1.15)	33 (1.56)	
	CTICU	85 (0.74)	18 (0.85)	
	Med-Surg ICU	7669 (66.71)	1345 (63.62)	
	MICU	1510 (13.14)	331 (15.66)	
	Neuro ICU	222 (1.93)	41 (1.94)	
	SICU	440 (3.83)	70 (3.31)	
weight (mean (SD))		83.271 (27.945)	79.428 (27.426)	<0.0001

Analysis platform

We utilize the widely used deep learning framework PyTorch (https://pytorch.org/) to construct our prediction model. PyTorch offers a comprehensive range of functionalities and empowers users with complete control over their Python programs [17]. Our prediction model is built on a Windows computer with the following specifications: Windows 11 operating system, 12th Gen Intel(R) Core (TM) i7-12700F central processing unit, 32 GB random access memory, and NVDIA GeForce RTX3060Ti graphics processing unit.

Model Architecture

The Transformer architecture, as illustrated in Figure 2, diverges from conventional CNN and RNN networks by primarily utilizing the Attention mechanism [9]. Specifically, the Transformer comprises self-Attention and Feed Forward Neural Network components exclusively. Considering the data format in our study, we have excluded the original embedding layer from the network

design. Instead, we have incorporated three layers of Encoder-Layer using the Transformer for extracting features from our data. Following a dropout layer, the extracted features are then fed into a fully connected network to obtain the prediction results. (Please refer to the following link for the relevant code: https://github.com/yanghaoljx/Sepsis-ICU-timeseries)

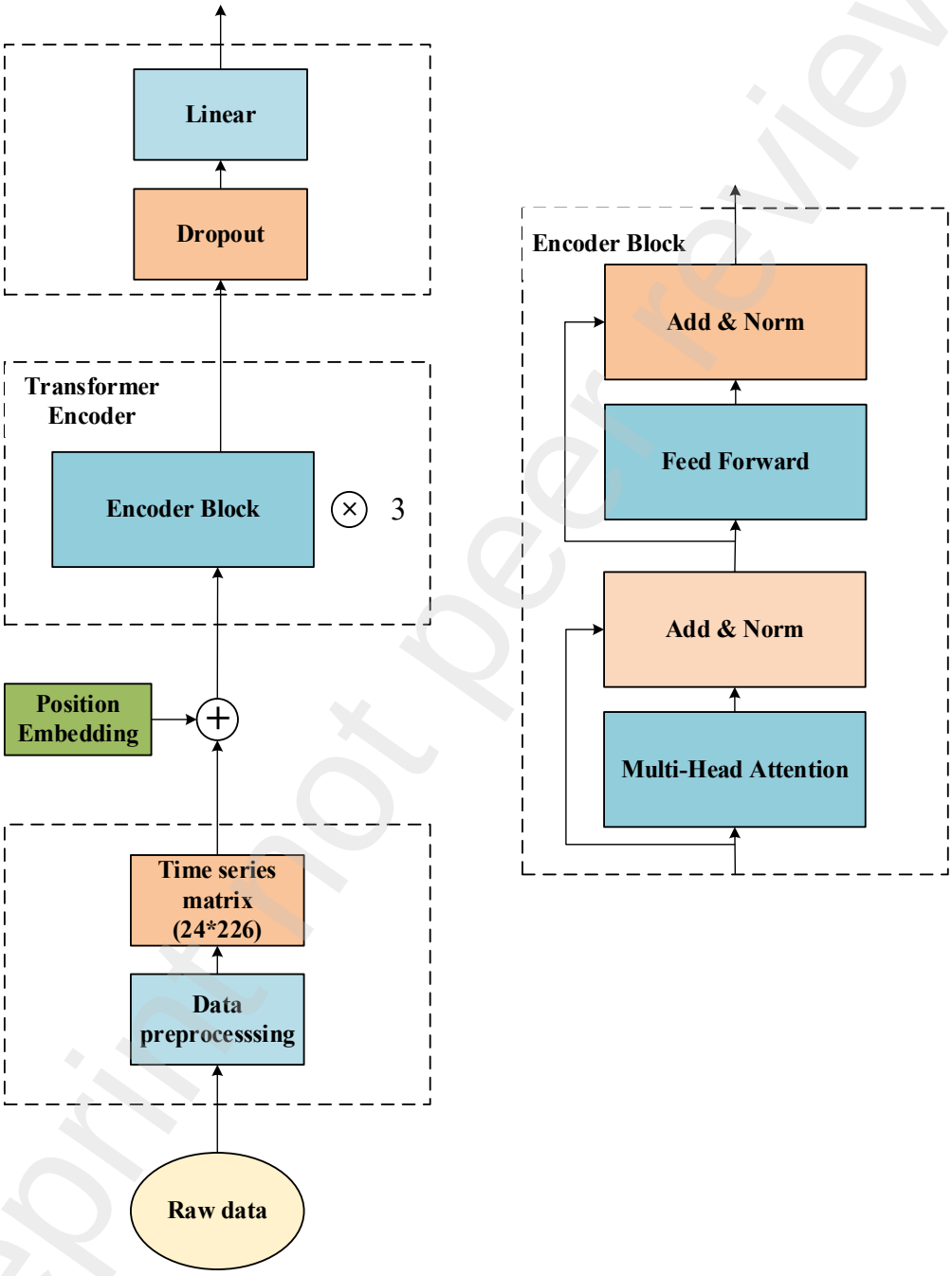

Figure 2. The architecture of Transformer

Utilizing the Transformer architecture, we develop network models for varying time periods based on the duration of ICU stay. Data features are collected for each hour since the patient's admission, forming a time-series model. These models are truncated according to the number of

days in the ICU, resulting in network models for different timeframes. The Encoder module within the Transformer is employed to extract data features. Subsequently, feature fusion is conducted on the models from different days, and the output results are obtained through a fully connected network. The overall framework is visualized in Figure 3, as depicted below.

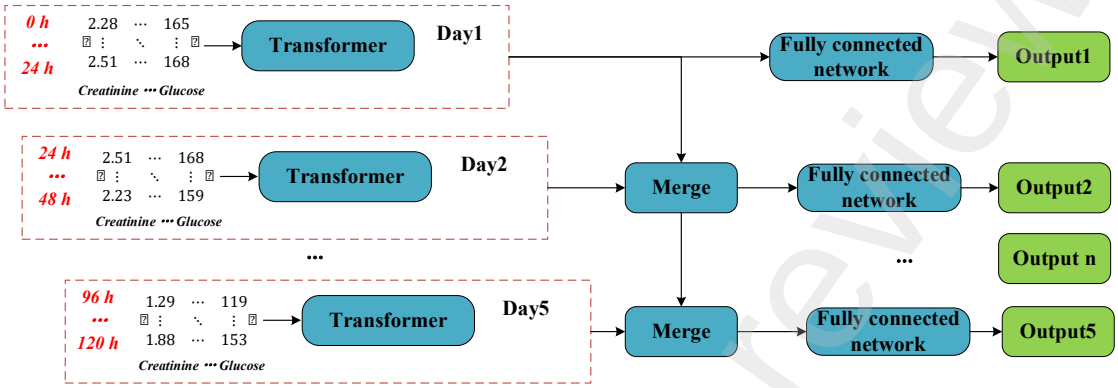

Figure 3. The framework of time-series model

Focal Loss

Considering the problem of imbalanced positive and negative samples in the study, the Focal loss [18] is introduced as the loss function for the model. It employs a dynamic scaling factor γ to dynamically reduce the weights of easily distinguishable samples during the training process, thereby quickly focusing on difficult-to-distinguish samples and reducing the contribution of easily classified samples to the loss. It also introduces weight factor $\alpha \in [0,1]$ to adjust the proportion between the losses of positive and negative samples.

$$FL(p_t) = -\alpha(1-p_t)^{\gamma}\log(p_t) \qquad (1)$$

SHAP

In 2017, Lundberg proposed the SHAP (Shapley additive explanations) method [19]. SHAP exhibits additivity consistency in explaining the output results, consistent with the general notion of regression. For each predicted sample, the model generates a predicted value, and the SHAP value is the assigned numerical value to each feature in that sample. Assuming the *i-th* sample is denoted as x_i, the *j-th* feature of the *i-th* sample is denoted as x_{ij}, the model's predicted value for the *i-th* sample is y_i, and the mean of all sample predictions is denoted as y_{base}. Then, the SHAP value of xij follows the following equation:

$$y_i = y_{base} + f(x_i,1) + f(x_i,2) + \ldots + f(x_i,k) \qquad (2)$$

Here, $f(x_i,j)$ represents the SHAP value of x_{ij}. This feature ensures that the sum of the

contribution values equals the final output, thereby eliminating the interpretability differences caused by structural variations among different models.

The main idea behind SHAP is derived from the Shapley value in cooperative game theory. Shapley developed this method to address the problem of allocating cooperative benefits among multiple players. In a set I={1,2,…,n} of n participants, if there exists a cooperative benefit function v for any subset S of I that satisfies $v(\emptyset)=0$ and for any disjoint subsets S_1 and S_2 of I exist $v(S_1 \cup S_2) \geq v(S_1) + v(S_2)$, then for each participant involved, the allocation of their contribution $x=\{x_1,x_2,\ldots,x_n\}$ must satisfy the following conditions for cooperation to occur: $\sum x_i = v(i)\ \ i = 1,2,...,n$, $x_i \geq v(i)$, i=1,2,…,n. In other words, each participant should receive no less than their individual contribution in a non-cooperative scenario, and the sum of the allocations should equal the total benefit. For the benefit function v, the allocation $\varphi(v) = (\varphi_1(v),\varphi_2(v),...,\varphi_n(v))$has been proven to satisfy the following properties:

$$\varphi_i(v) = \sum_{i \in T}^{T \in I} \frac{(|T|-1)!(n-|T|)!}{n!}(v(T) - v(T-\{i\})) \qquad (3)$$

Here, T is a subset of the set, $|T|$ represents the number of elements in the subset, and n represents the total number of members. SHAP builds upon the Shapley value and makes improvements suitable for Machine learning models. It treats features as players and model outputs as the cooperative results. For the *K-th* data point, the model output can be represented as $v_k(I)$, The contribution value of each feature at that data point, denoted as $\varphi_i(v_k)$, is also known as the SHAP value. Unlike linear models that use the magnitude of parameters or coefficients to measure the contribution of a feature to the model, the SHAP algorithm calculates the combined contributions of each feature for every sample. In the end, we obtain the contribution of each feature in each sample. If a feature exhibits consistent trends across the majority of samples, it indicates that the model recognizes the importance of that feature in either a positive or negative direction.

Results

Based on the temporal features of patients on each day, a Transformer network model is built to obtain their prediction results. The network model for the *N-th* day incorporates the encoder features of the previous *N-1* days. It is observed that the model performs the best in predicting the results for the 5th day. The following comparison shows the daily prediction results in comparison

with those obtained using the traditional Apache II-based prediction method.

Table 1.The outcomes of prediction models

Model	Day1	Day2	Day3	Day4	Day5
AUC	0.87(0.021)	0.88(0.022)	0.89(0.019)	0.92(0.011)	0.92(0.009)
Accuracy	77.82(1.63)	78.93(1.53)	80.98 (0.89)	82.03 (0.71)	83.92(0.69)
F1-score	0.704 (0.058)	0.711(0.061)	0.732(0.044)	0.773(0.048)	0.804(0.024)

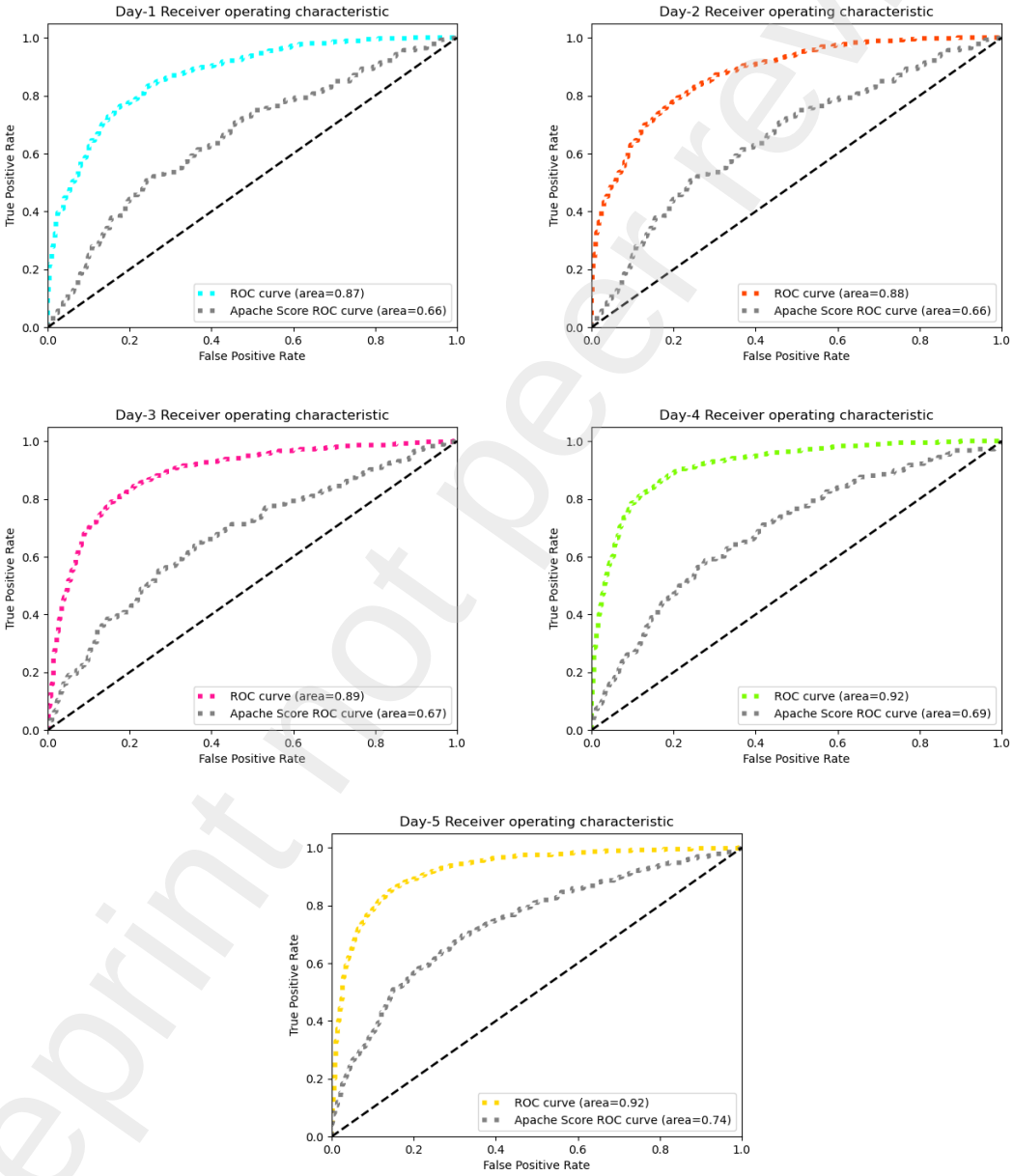

Figure 4. ROC curves for different days in the model.

Utilizing the prediction model, weight analysis is conducted using the SHAP algorithm, which generates visual heatmaps. Due to limitations in formatting, the paper presents only the

heatmap representing the most prominently activated feature for each day, as depicted below.

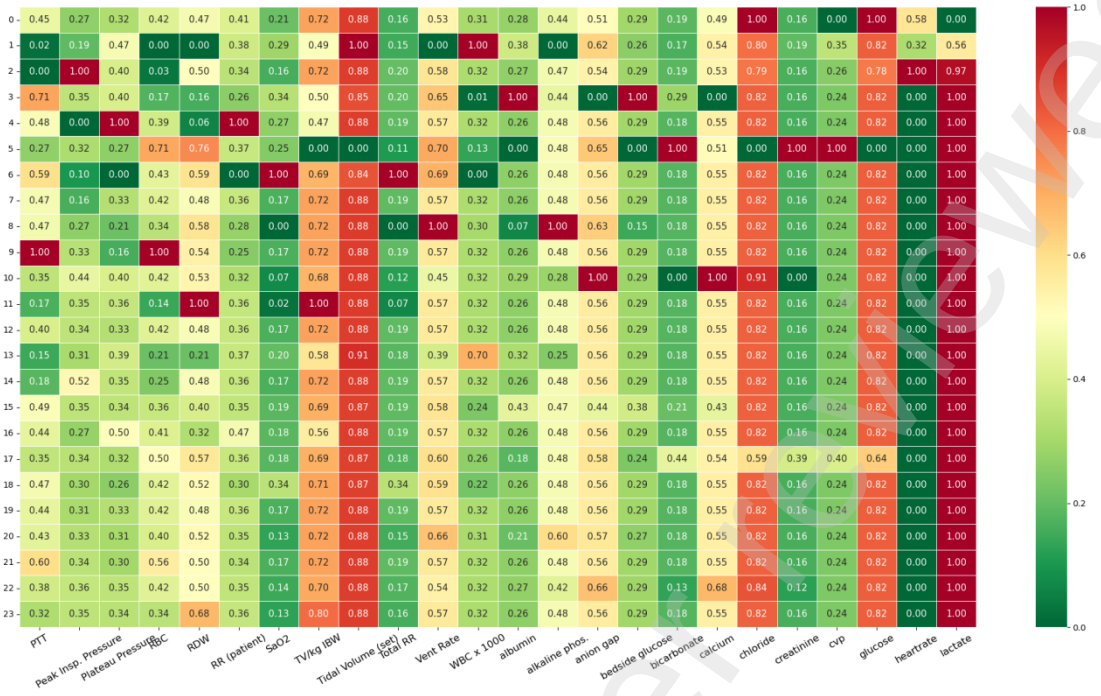

Figure 5. Day-1 model visualization results, with prominently activated features including Lactate, tidal volume, chloride, glucose

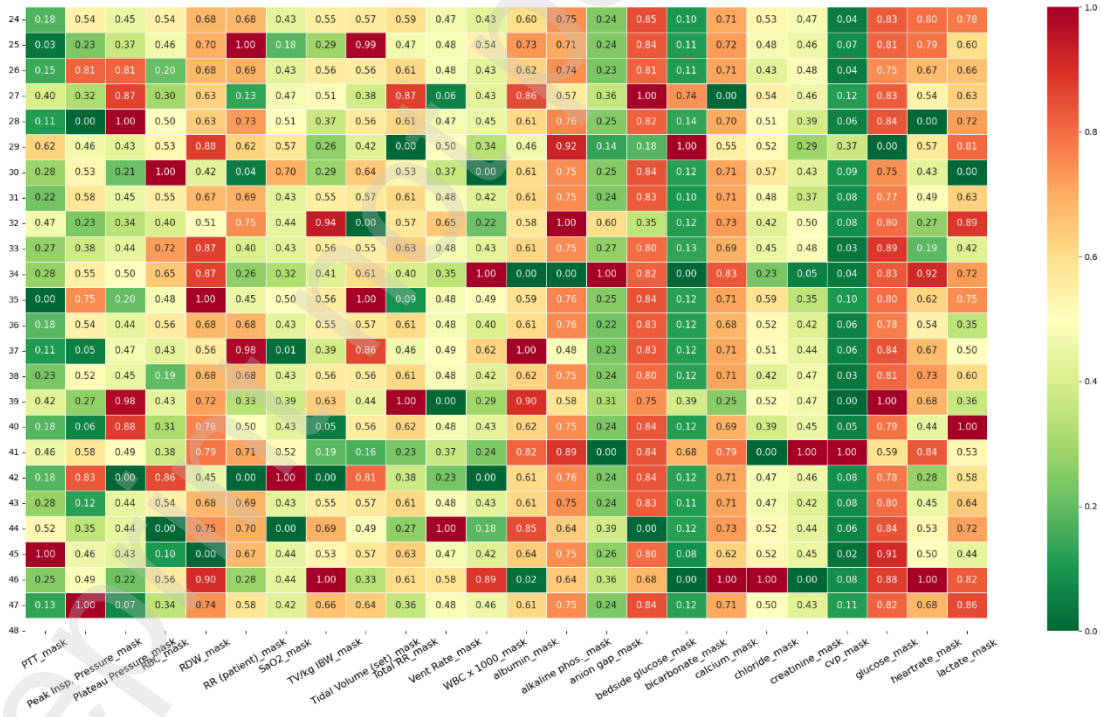

Figure 6. Day-2 model visualization results, with prominently activated features including:

RDW, albumin, alkaline, glucose, calcium

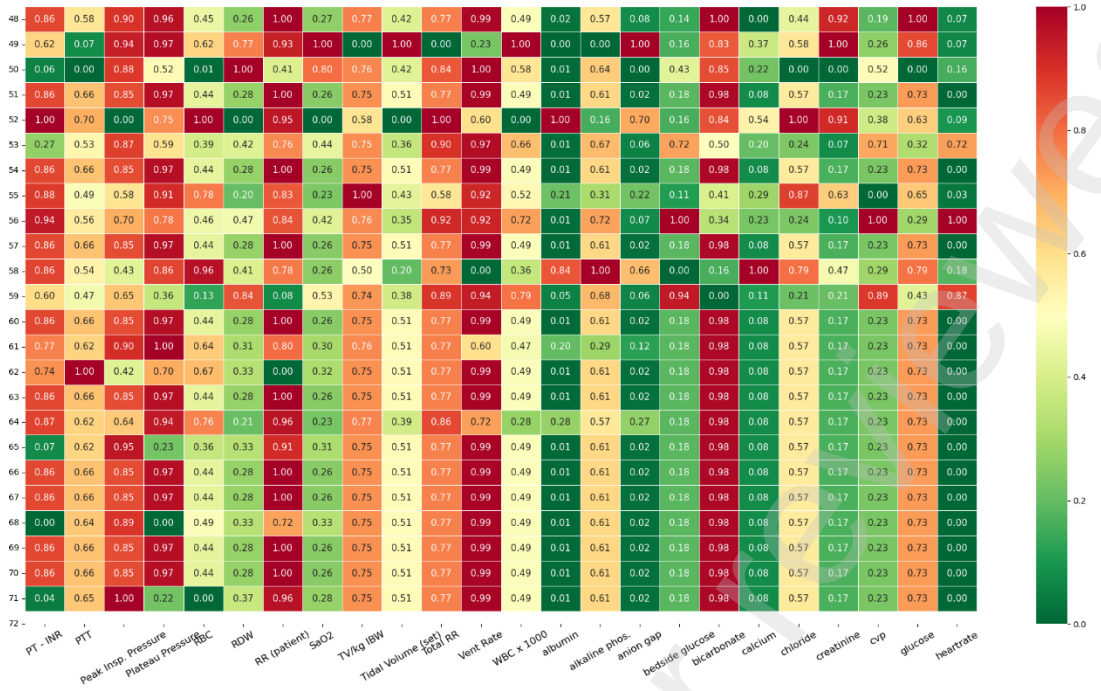

Figure 7. Day-3 model visualization results, with prominently activated features including:

PT-INR, PTT, Peak Insp.Pressure, plateau pressure, Total RR, TV/kg IBW, Total RR, Vent Rate, bicarbonate

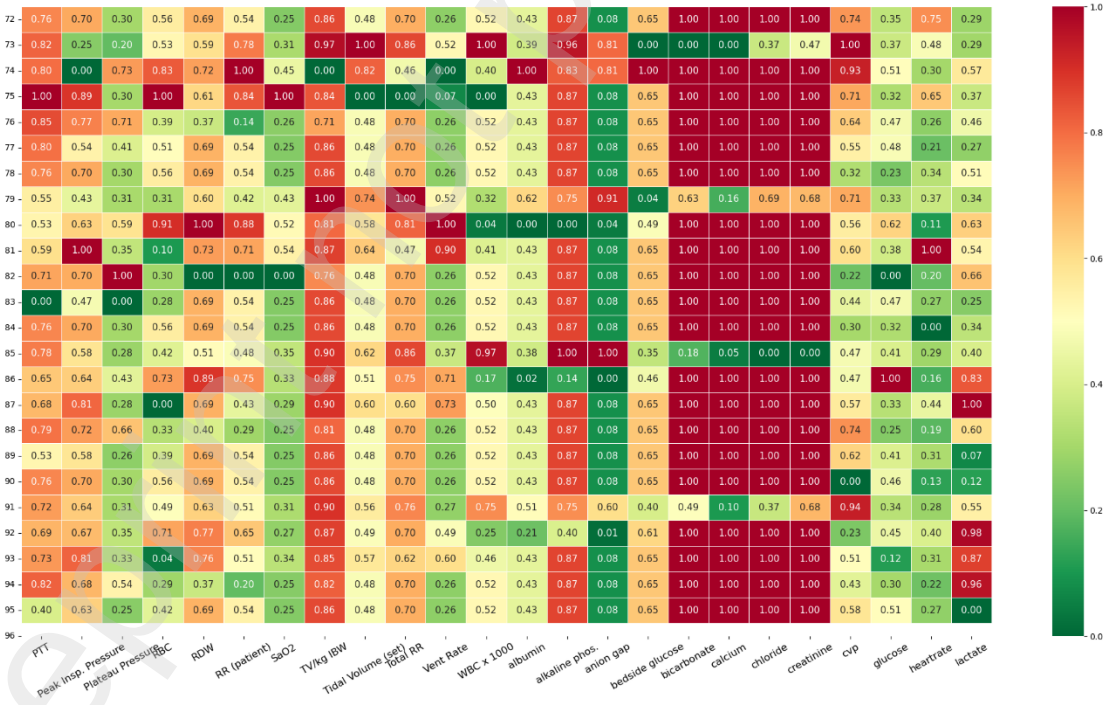

Figure 8. Day-4 model visualization results, with prominently activated features including:

PTT, Peak Insp.Pressure, TV/kg IBW, Total RR, creatine, chloride, calcium, bicarbonate

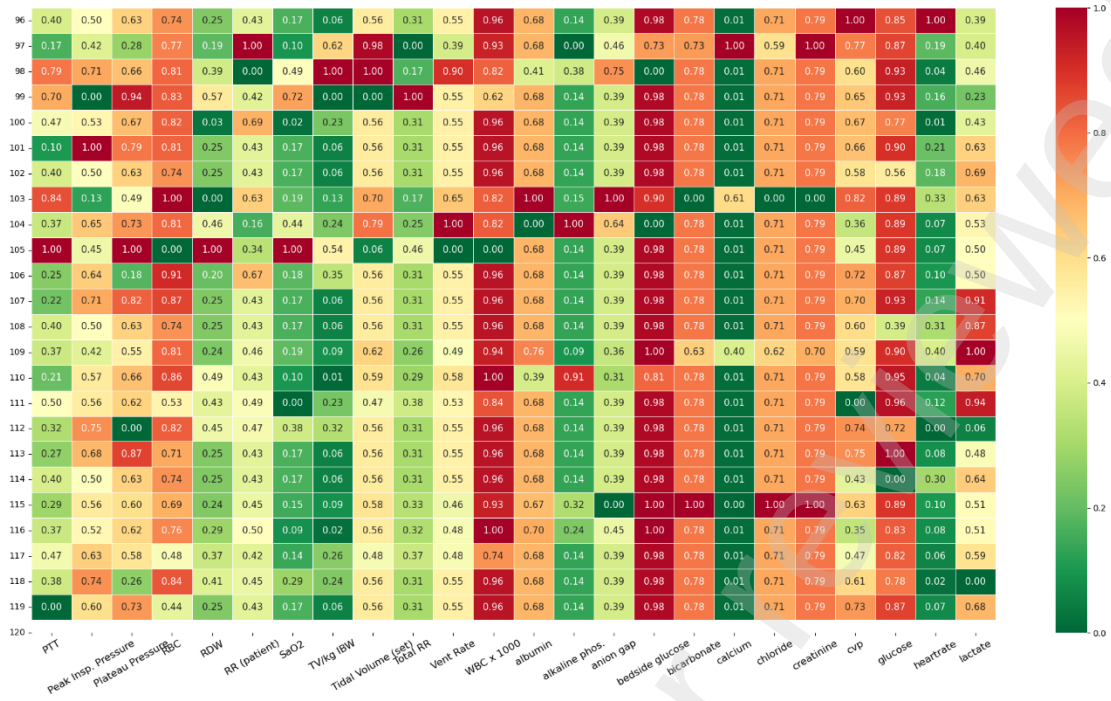

Figure 9. Day-5 model visualization results, with prominently activated features including:

RBC, WBC, glucose,bicarbonate,chloride,creatinine

Discussion

Sepsis constitutes a perilous medical state demanding the precise of patients at elevated risk to optimize therapeutic interventions and prognostic projections. In this study, we propose a time-series modeling approach based on Transformers to predict the outcomes of patients in the ICU

The findings from our model indicate that the Transformer-based time-series model exhibits a gradual improvement in predictive accuracy as the duration of patients' stay in the ICU increases. Through receiver operating characteristic (ROC) Curve, our model showed an AUC from 0.87 (±0.021) on the first day to 0.92 (±0.009) on the fifth day. The observed trend pointed out that incorporating a broader range of patients' features indeed enhanced the reliability of the model. Furthermore, by employing the SHAP visualization algorithm, we noted a rising presence of life features strongly linked to mortality as the time sequence advances. This is distinctly evident from their elevated levels of activation. Comparing these obtained activation features with relevant literature, we found that it has certain clinical and research value.

The vital signs and diagnostic indicators recorded during the patient's initial day of admission must not be overlooked, as they play a pivotal role in evaluating the severity of the patient's condition. The model on the first day, lactate levels was the most strongly associated features with

patient mortality after admission to the ICU, then followed by tidal volume, chloride, bedside glucose, and IBW. In most cases, elevated lactate levels were correlated with poor prognosis, particularly in sepsis, trauma, bleeding, shock, and cardiac arrest [20-22].Lactate levels possess diverse diagnostic applications and have been extensively utilized as an indicator for resuscitation, the stratification of risk, and the prediction of mortality in cases of sepsis. [23]. In a prospective cohort study of patients with infections in the emergency department (ED), elevated lactate levels were associated with increased mortality, and an initial lactate level equal to or greater than 4 mmol/L was found to be linked to a 28% in-hospital mortality rate [24]. ICU patients who require mechanical ventilation face a high incidence and mortality rate [25-26]. Acute respiratory distress syndrome (ARDS) contributes to the elevated admission and mortality rates in the ICU, and lung-protective ventilation, including low tidal volume ventilation, has demonstrated benefits for both ARDS and non-ARDS patients [27-28]. According to our model, in addition to lactate concentration, electrolyte disturbances are also common in the ICU. Serum chloride, the second most abundant electrolyte in the human body, plays a critical role in the pathophysiology of acute heart failure and is considered a therapeutic target for reducing mortality [29-31]. Among ICU patients with heart failure, various electrolyte abnormalities were commonly observed, with hyponatremia and hypochloremia being the most prevalent [32]. Previous studies have investigated the detrimental effects of hyperglycemia on the prognosis of critically ill patients, such as those with sepsis, myocardial infarction, acute pancreatitis, and stroke [33]. Our model enables us not only to observe the indicators corrected with human organ failure, but also the systemic indicators like glucose levels. In our research, serum blood glucose concentration during ICU stay was associated with overall ICU mortality in sepsis patients. Patients with severe hyperglycemia ($\geq$200 mg/dL) upon ICU admission exhibit the highest overall ICU mortality, regardless of whether they have diabetes, highlighting sustained hyperglycemia as a significant risk factor for ICU mortality in sepsis patients [34].

As the disease advances and in response to the subsequent administration of relevant medications, surgical interventions, and other treatment modalities, the patient's indicators will undergo alterations. The model on the second day, the levels of RDW (Red Cell Dist ribution Width), albumin, alkaline phosphatase, and calcium start to increase. Following res

piratory support and basic care, most patients show signs of recovery. However, patients with organ failure begin to exhibit disruptions in relevant organ function indicators. RDW is a parameter traditionally used for the differential diagnosis of anemia, and its increase r eflects disruptions in red blood cell production and unregulated red cell homeostasis. Tutak et al. conducted a study indicating that elevated RDW levels serve as a reliable marker fo r predicting mortality and should be included in the APACHE II score to forecast patient outcomes [35]. RDW values above the upper limit of 14.5% are associated with abnormal metabolic conditions such as inflammation, oxidative stress, and nutritional imbalances [36]. Özdemir et al. demonstrated an association between high RDW levels and mortality in IC U patients with mixed conditions [37]. Similarly, Safdar et al. reported a correlation betwe en high RDW levels and 30-day mortality rates in follow-up patients from ICU medical a nd surgical departments [38]. Ochoa et al. also found a correlation between elevated RDW levels and mortality in ICU patients [39]. From this, it can be seen that the Red Cell Dis tribution Width (RDW) in routine blood tests plays a significant role in predicting the mo rtality rate and prognosis of ICU sepsis patients, and should not be overlooked. Moreover, our model showed another indicator associated with sepsis severity. Albumin, a multifuncti onal protein with colloidal and pharmacological properties, exhibits physiological functions intricately linked to its distribution (intravascular, extravascular, and intracellular locations), concentration, and complex structure. [40]. Yin et al. conducted a prospective cohort study and found that serum albumin levels below 2.92 g/dL upon admission are associated with an increased 28-day mortality rate in patients with severe sepsis [41]. Research suggested that serum albumin trends can predict mortality rates in ICU sepsis patients. Significant co rrelations exist between mortality rates and serum albumin trends, as well as average, peak, admission, and minimum levels of albumin [42]. In critically ill elderly patients, serum alb umin levels upon ICU admission serve as effective predictive indicators for mortality and other outcomes. Low serum albumin levels upon admission were independent risk factors f or six-month mortality in critically ill elderly patients after ICU discharge [43]. Sepsis is associated with organ damage. Acute kidney injury (AKI) is one of the most common org an failures. Total alkaline phosphatase activity slightly increases in ICU patients with septi c AKI, and its activity positively correlates with ICU length of stay [44]. In a predictive

experiment for early liver failure in pediatric ICU patients, elevated alkaline phosphatase o n the first day of admission was identified as a key laboratory parameter for diagnosing li ver dysfunction. Alkaline phosphatase is widely used to predict liver function prognosis in ICU patients [45].

The parameters most strongly associated with ICU mortality were identified, with one-third of sepsis-related deaths occurring within three days of ICU admission [46]. Therefore, in the models on the third and fourth days, coagulation parameters such as prothrombin time (PT) and activated partial thromboplastin time (APTT) reflect their value in predicting the length of hospital stay for septic patients. PT and APTT can be served as early prognostic markers for severe pneumonia requiring transfer to the ICU. Denis et al. compared these coagulation parameters and various coagulation and inflammation factor levels between patients with early treatment-resistant respiratory failure (RRF) or severe acute respiratory distress syndrome (SARDS) requiring ICU treatment and stable COVID-19 patients. They discovered that early measurement of PT values can predict the course of COVID-19-related pneumonia. Prolonged PT in patients transferred to the ICU and PT levels upon admission for COVID-19 patients can serve as early prognostic markers for severe pneumonia requiring ICU care [47-48]. APTT, the most commonly used and sensitive screening method for assessing clotting activity, has been shown in multiple studies to be a risk factor for thrombosis and pulmonary embolism. It is a significant predictor for pulmonary embolism in ICU patients [49-52]. Monitoring APTT in ICU patients should be emphasized to prevent pulmonary embolism. For patients admitted to the ICU due to respiratory conditions, particularly COVID-19, respiratory-related parameters exhibit a strong association with ICU mortality. Plateau pressure, which represents airway pressure within the alveoli during positive pressure ventilation when the breath is held, has been analyzed for its impact on the prognosis of patients with acute respiratory distress syndrome (ARDS). There is a correlation between plateau pressure and short-term mortality, with significantly higher overall short-term mortality rates observed in patients with plateau pressure exceeding 32 cm H2O within three days of ICU admission. This correlation becomes evident in the following days and may persist for the first three days after ICU admission [53]. The respiratory index (RI), the ratio of peripheral oxygen saturation (SpO2) to respiratory rate (RR), has been validated for prognostic stratification in patients with acute pulmonary embolism[54]. However, in patients admitted to the ICU due to COVID-19, RR may be an important parameter

that has received insufficient reporting and attention. Giorgio et al. found that RR is a better predictor of mortality compared to blood gas parameters (P/F and STP/F) in evaluating patients with COVID-19-related pneumonia admitted to the ICU [55].

In the model on the fifth day, compared to the previous four days, red blood cell, white blood cell, and creatinine levels began to rise. Cells of the innate immune system and adaptive immune system play a crucial role in the host response to sepsis, which is associated with profound inhibition of constitutive neutrophil apoptosis[56]. During the onset of sepsis, neutrophils respond rapidly to infection, their numbers increase sharply, and they quickly migrate to sites of severe infection[57]. While neutrophil counts are indicative of the overall inflammation severity, in intricate sepsis scenarios, the delay in neutrophil apoptosis could result in a persistent elevation of neutrophil counts.[58]. Hence, the significance of leukocytes as an indicator should not be diminished in the presence of sepsis; rather, it should continue to be closely monitored. The incidence of acute kidney injury (AKI) in sepsis patients in the intensive care unit (ICU) can reach up to 60%, and the occurrence of AKI is associated with increased mortality, with survivors being at risk of developing chronic kidney disease[59-60]. Serum creatinine slowly increases after the onset of AKI, and its rise can be further delayed due to large fluid resuscitation and fluid balance[61]. On the third day following sepsis-induced AKI, serum creatinine may only increase to the borderline value. If glomerular filtration rate remains stable, the expansion of distribution volume is expected to lead to a decrease in serum creatinine levels, resulting in a delay in the number of days for serum creatinine concentration to rise [62]. Therefore, continuous monitoring of creatinine levels in sepsis patients in the ICU allows for better assessment of renal injury.

Conclusion

The results of this study highlight the effectiveness of time-series algorithms, specifically the Transformer-based approach, in predicting patient outcomes in the ICU. The utilization of these algorithms enables the capture of temporal patterns and the integration of comprehensive patient features, leading to improved accuracy and reliability in predictions. Furthermore, the identification and activation of life features associated with mortality provide valuable insights into the underlying mechanisms of sepsis and its progression.

In future research, it would be beneficial to explore the clinical implications of these identified activation features and examine their potential role in guiding early interventions and personalized

treatment plans for sepsis patients. Additionally, advancing the development of more sophisticated time-series models by incorporating additional data sources like MetaSepsisBase [63] and refining feature selection techniques holds promise for further enhancing the predictive capabilities and practical applications of mortality prediction in critical care settings.

Reference

1. Centers for Disease Control and Prevention. What is sepsis? 2021.
2. National Center for Health Statistics. About multiple cause of death, 1999–2019. CDC WONDER online database. 2020.
3. Lambden S, Laterre P F, Levy M M, et al. The SOFA score—development, utility and challenges of accurate assessment in clinical trials[J]. Critical Care, 2019, 23(1): 1-9.
4. Zeng W, Mao P, Huang Y, et al. Analyses of factors affecting prognosis of patients with sepsis and evaluation of their predicting values[J]. Chinese Journal of Integrated Traditional and Western Medicine in Intensive and Critical Care, 2015: 118-123.
5. Rhee C, Wang R, Song Y, Zhang Z, Kadri SS, Septimus EJ, et al. Risk adjustment for sepsis mortality to facilitate hospital comparisons using centers for disease control and prevention's adult sepsis event criteria and routine electronic clinical data. Crit Care Explor 2019;1(10):e0049
6. Park J Y, Hsu T C, Hu J R, et al. Predicting sepsis mortality in a population-based national database: Machine learning approach[J]. Journal of Medical Internet Research, 2022, 24(4): e29982.
7. Su L, Xu Z, Chang F, et al. Early prediction of mortality, severity, and length of stay in the intensive care unit of sepsis patients based on sepsis 3.0 by machine learning models[J]. Frontiers in Medicine, 2021, 8: 664966.
8. Pollard T J, Johnson A E W, Raffa J D, et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research[J]. Scientific data, 2018, 5(1): 1-13.
9. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
10. Oliphant T E. A guide to NumPy[M]. USA: Trelgol Publishing, 2006.
11. Reback J, McKinney W, Van Den Bossche J, et al. pandas-dev/pandas: Pandas 1.0. 5[J]. Zenodo, 2020.
12. Pantanowitz A, Marwala T. Missing data imputation through the use of the random forest algorithm[C]//Advances in computational intelligence. Springer Berlin Heidelberg, 2009: 53-62.
13. PostgreSQL B. PostgreSQL[J]. Web resource: http://www. PostgreSQL. org/about, 1996.

14. Pramesti G. Kupas tuntas data penelitian dengan SPSS 22[M]. Elex Media Komputindo, 2015.
15. McKnight P E, Najab J. Mann-Whitney U Test[J]. The Corsini encyclopedia of psychology, 2010: 1-1.
16. McHugh M L. The chi-square test of independence[J]. Biochemia medica, 2013, 23(2): 143-149.
17. Imambi S, Prakash K B, Kanagachidambaresan G R. PyTorch[J]. Programming with TensorFlow: Solution for Edge Computing Applications, 2021: 87-104.
18. Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.
19. Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Advances in neural information processing systems, 2017, 30.
20. Donnino MW, Miller J, Goyal N, et al. Effective lactate clearance is associated with improved outcome in post-cardiac arrest patients. Resuscitation. 2007; 75:229-234.
21. Regnier MA, Raux M, Le Manach Y, et al. Prognostic significance of blood lactate and lactate clearance in trauma patients. Anesthesiology. 2012; 117:1276-1288.
22. Wardi G, Wali AR, Villar J, et al. Unexpected intensive care transfer of admitted patients with severe sepsis. J Intensive Care.2017;5:43.
23. Demystifying Lactate in the Emergency Department
24. Shapiro NI, Howell MD, Talmor D, et al. Serum lactate as a predictor of mortality in emergency department patients with infection. Ann Emerg Med. 2005; 45:524-528.
25. Rubenfeld G.D., Herridge M.S. Epidemiology and outcomes of acute lung injury. Chest. 2007; 131:554–562. [PubMed] [Google Scholar]
26. Rose L., Gray S., Burns K., et al. Emergency department length of stay for patients requiring mechanical ventilation: a prospective observational study. Scand J Trauma Resusc Emerg Med. 2012; 20:30. [PMC free article] [PubMed] [Google Scholar]
27. Acute Respiratory Distress Syndrome: Challenge for Diagnosis and Therapy
28. Neto A.S., Cardoso S.O., Manetta J.A., et al. Association between use of lung-protective ventilation with lower tidal volumes and clinical outcomes among patients without acute respiratory distress syndrome: a meta-analysis. JAMA. 2012; 308:1651–1659. [PubMed]

[Google Scholar]

29. Waikar SS, Mount DB, Curhan GC. Mortality after hospitalization with mild, moderate, and severe hyponatremia. Am J Med. 2009 Sep;122(9):857–65.
30. Berend K, van Hulsteijn LH, Gans RO. Chloride: the queen of electrolytes? Eur J Intern Med. 2012 Apr 1;23(3):203–11.
31. Grodin JL, Simon J, Hachamovitch R, Wu Y, Jackson G, Halkar M, et al. Prognostic role of serum chloride levels in acute decompensated heart failure. J Am Coll Cardiol. 2015 Aug;66(6):659–66
32. The Role of Serum Chloride in Acute and Chronic Heart Failure: A Narrative Review
33. Hou D, Zhong P, Ye X, Wu D. Persistent hyperglycemia is a useful glycemic pattern to predict stroke mortality: a systematic review and meta-analysis. BMC Neurol. (2021) 21:487. 10.1186/s12883-021-02512-1 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
34. Lu Z, Tao G, Sun X, Zhang Y, Jiang M, Liu Y, Ling M, Zhang J, Xiao W, Hua T, Zhu H, Yang M. Association of Blood Glucose Level and Glycemic Variability With Mortality in Sepsis Patients During ICU Hospitalization. Front Public Health. 2022 Apr 29; 10:857368. doi: 10.3389/fpubh.2022.857368. PMID: 35570924; PMCID: PMC9099235.
35. Tutak AS, Avni Findikli H. Can RDW be Added to Intensive Care Disease Severity Scoring Systems? Arch Clin Biomed Res 2018; 2: 183-18
36. Fernandez R, Cano S, Catalan I, Rubio O, Subira C, Masclans J, Rognoni G, Ventura L, Macharete C, Winfield L, Alcoverro JM. High red blood cell distribution width as a marker of hospital mortality after ICU discharge: a cohort study. J Intensive Care 2018; 6: 1-7
37. Özdemir R, Mutlu NM, Özdemir M, Akçay M, Yel C, Turan IÖ. The importance of red cell distribution width (RDW) in patient follow up in intensive care unit (ICU). Acta Medica Mediterr 2016; 32: 349-354.
38. Safdar S, Modi T, Sriramulu L, Shaaban H, Sison R, Modi V, Adelman M, Guron G. The role of red cell distribution width as a predictor of mortality for critically ill patients in an inner-city hospital. J Nat Sci Biol Med 2017; 8: 154.
39. Ochoa SH, Martínez MI, Carrillo RSC, Esponda PJG. Extent of erythrocyte distribution as a predictor of mortality on admission to the Intensive Care Unit of the Angeles Pedregal Hospital. Acta Med 2019; 17: 230-236.

40. Garcia-Martinez R, Caraceni P, et al. Hepatology, 2013,58:1836-1846.

41. Yin, M., Si, L., Qin, W., Li, C., Zhang, J., Yang, H., ... Wang, H. (2016). Predictive value of serum albumin level for the prognosisof severe sepsis without exogenous human albumin administration:A prospective cohort study. Journal of Intensive Care Medicine,33, 687–694. doi:10.1177/0885066616685300

42. Pan SW, Kao HK, Yu WK, Lien TC, Chen YW, Wang JH, Kou YR. Synergistic impact of low serum albumin on intensive care unit admission and high blood urea nitrogen during intensive care unit stay on post-intensive care unit mortality in critically ill elderly patients requiring mechanical ventilation. Geriatr Gerontol Int. 2013 Jan;13(1):107-15. doi: 10.1111/j.1447-0594.2012.00869.x. Epub 2012 Jun 4. PMID: 22672187.

43. Arques S, Roux E, Stolidi P, Gelisse R, Ambrosi P. Usefulness of serum albumin and serum total cholesterol in the prediction of hospital death in older patients with severe, acute heart failure. Arch Cardiovasc Dis 2011; 104:502–508.

44. Baek SD, Kang JY, Yu H, Shin S, Park HS, Kim MS, Lee EK, Kim SM, Chang JW. Change in alkaline phosphatase activity associated with intensive care unit and hospital length of stay in patients with septic acute kidney injury on continuous renal replacement therapy. BMC Nephrol. 2018 Sep 20;19(1):243. doi: 10.1186/s12882-018-1028-9. PMID: 30236070; PMCID: PMC6148963.

45. Zahmatkeshan M, Serati Z, Freydooni S, Safarpour AR, Esmailnejad A, Haghbin S. Prediction of Early Liver Failure in Pediatric Patients Admitted to Intensive Care Unit. Middle East J Dig Dis. 2019 Jul;11(3):141-146. doi: 10.15171/mejdd.2019.140. Epub 2019 Jun 14. PMID: 31687112; PMCID: PMC6819968.

46. Daviaud F, Grimaldi D, Dechartres A, et al. Timing and causes of death in septic shock. Ann Intensive Care. 2015;5(1):16. doi:10.1186/s13613-015-0058-8

47. Baranovskii DS, Klabukov ID, Krasilnikova OA, Nikogosov DA, Polekhina NV, Baranovskaia DR, Laberko LA. Prolonged prothrombin time as an early prognostic indicator of severe acute respiratory distress syndrome in patients with COVID-19 related pneumonia. Curr Med Res Opin. 2021 Jan;37(1):21-25. doi: 10.1080/03007995.2020.1853510. Epub 2020 Dec 7. PMID: 33210948; PMCID: PMC7738209.

48. Tang N, Li D, Wang X, et al. Abnormal coagulation parameters are associated with poor

prognosis in patients with novel corona virus pneumonia. J Thromb Haemost. 2020;18(4):844–847.

49. Prucnal CK, et al. Analysis of partial thromboplastin times in patients with pulmonary embolism during the first 48 hours of anticoagulation with unfractionated heparin. Acad Emerg Med. 2020;27(2):117-127. [PubMed] [Google Scholar]

50. Lippi G, Favaloro EJ, Franchini M. Paradoxical thrombosis part 1: factor replacement therapy, inherited clotting factor deficiencies and prolonged APTT. J Thromb Thrombolysis. 2012;34(3):360-366. [PubMed] [Google Scholar]

51. Matsubara S, et al. Prolonged activated partial thromboplastin time in thromboprophylaxis with unfractionated heparin in patients undergoing cesarean section. J Obstet Gynaecol Res. 2010;36(1):58-63. [PubMed] [Google Scholar]

52. Huang CB, Hong CX, Xu TH, Zhao DY, Wu ZY, Chen L, Xie J, Jin C, Wang BZ, Yang L. Risk Factors for Pulmonary Embolism in ICU Patients: A Retrospective Cohort Study from the MIMIC-III Database. Clin Appl Thromb Hemost. 2022 Jan-Dec;28:10760296211073925. doi: 10.1177/10760296211073925. PMID: 35043708; PMCID: PMC8796081.

53. Yasuda H, Sanui M, Nishimura T, Kamo T, Nango E, Abe T, Roberts R, Takebayashi T, Hashimoto S, Lefor AK. Optimal Upper Limits of Plateau Pressure for Patients With Acute Respiratory Distress Syndrome During the First Seven Days: A Meta-Regression Analysis. J Clin Med Res. 2021 Jan;13(1):48-63. doi: 10.14740/jocmr4390. Epub 2021 Jan 12. PMID: 33613800; PMCID: PMC7869567.

54. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A. Respiratory rate: the neglected vital sign. Med J Aust. 2008;188(11):657–9.

55. Maraziti G, Becattini C. Early Variation of Respiratory Indexes to Predict Death or ICU Admission in Severe Acute Respiratory Syndrome Coronavirus-2-Related Respiratory Failure. Respiration. 2022;101(7):632-637. doi: 10.1159/000522275. Epub 2022 Mar 15. PMID: 35290981; PMCID: PMC9059089.

56. Taneja R, Parodo J, Kapus A, Rotstein OD, Marshall JC: Delayed neutrophil apoptosis in sepsis is associated with maintenance of mitochondrial transmembrane potential (DYM) and reduced caspase-9 activity. Crit Care Med 32(7):1460–1469, 2004.

57. Huang Z, Fu Z, Huang W, Huang K. Prognostic value of neutrophil-to-lymphocyte ratio in sepsis: A meta-analysis. Am J Emerg Med. 2020 Mar;38(3):641-647. doi: 10.1016/j.ajem.2019.10.023. Epub 2019 Nov 18. PMID: 31785981.

58. Taneja R, Parodo J, Jia SH, Kapus A, Rotstein OD, Marshall JC. Delayed neutrophil apoptosis in sepsis is associated with maintenance of mitochondrial transmembrane potential and reduced caspase-9 activity. Crit Care Med 2004;32(7):1460-9.

59. Lameire NH, Bagga A, Cruz D, De Maeseneer J, Endre Z, Kellum JA, Liu KD, Mehta RL, Pannu N, Van Biesen W, Vanholder R. Acute kidney injury: an increasing global concern. Lancet. 2013 Jul 13;382(9887):170-9. doi: 10.1016/S0140-6736(13)60647-9. Epub 2013 May 31. PMID: 23727171.

60. Hoste EA, Bagshaw SM, Bellomo R, et al. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. Intensive Care Med. 2015;41(8):1411-1423. doi:10.1007/s00134-015-3934-7

61. Pickering JW, Ralib AM, Endre ZH. Combining creatinine and volume kinetics identifies missed cases of acute kidney injury following cardiac arrest. Crit Care. 2013;17(1):R7. Published 2013 Jan 17. doi:10.1186/cc11931

62. Pickkers P, Mehta RL, Murray PT, et al. Effect of Human Recombinant Alkaline Phosphatase on 7-Day Creatinine Clearance in Patients With Sepsis-Associated Acute Kidney Injury: A Randomized Clinical Trial. JAMA. 2018;320(19):1998-2009. doi:10.1001/jama.2018.14283

63. Zhang C, Zhang X, Sun Z, Liu X, Shen B. MetaSepsisBase: a biomarker database for systems biological analysis and personalized diagnosis of heterogeneous human sepsis. Intensive Care Med. 2023 Jun 17. doi: 10.1007/s00134-023-07126-4. Epub ahead of print. PMID: 37329364.