DOI: 10.4149/BLL_2023_037

Bratisl Med J 2023; 124 (3)
239 – 250

ANALYTICAL STUDY

Detection of sepsis using biomarkers based on machine learning

Mahsa Amin ESKANDARI [1], Mohammad Karimi MORIDANI[2], Salar MOHAMMADI [1]

Department of Biomedical Engineering, Faculty of Health, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran. **karimi.m@iautmu.ac.ir**

ABSTRACT
BACKGROUND: Sepsis is the second most common cause of death in patients with non-cardiovascular diseases admitted to the ICU. It is one of the top ten reasons for death among all hospitalized patients. This study aimed to compare the value of some blood parameters in diagnosing sepsis and investigate their relationship to select a more practical diagnostic method.
METHODS: In this descriptive-analytical study, 208 patients with sepsis admitted to the ICU were selected. Then the physiological parameters of patients and normal individuals were measured. Data analysis was performed using the p value and effect size methods and MATLAB software. To classify the disease, the MLP, RBF, and KNN methods were used.
RESULTS: The values of the HR, O_2Sat, and SBP in patients with sepsis have changed significantly compared to NORMAL conditions. The classification results using different classifications showed that the values of specificity, sensitivity, and accuracy values in the classifier are more than MLP and RBF and equal to 98 %, 100 %, and 99 %, respectively.
CONCLUSIONS: Clinically, accurate detection of sepsis and predicting the patients at risk of developing sepsis is useful for improving treatment. Given the significant differences between HR, O_2Sat, and SBP between normal and sepsis patients in this study, it may be possible to use these tests as simple tests instead of the complement protein 3 (C3) and Procalcitonin (PCT) tests to diagnose sepsis in the ICU *(Tab. 8, Fig. 10, Ref. 39)*. Text in PDF *www.elis.sk*
KEY WORDS: sepsis, physiological parameters, detection; feature extraction, statistical analysis.

Abbreviations: ANOVA – Analysis of variance, APACHE – Acute Physiology and Chronic Health Evaluation, CA – Cardiac Arrest, EGDT – Early Goal-Directed Therapy, HR – Heart Rate, ICU – Intensive Care Unit, KNN – K-Nearest Neighbor, MLP – Multi-Layer perceptron, PCT – Procalcitonin, PSV – Pipe Separated Value, RBF – Radial Basic Function, ROSC – Return of spontaneous circulation, SAPS – Simplified Acute Physiology Score, SBP – Spontaneous Bacterial Peritonitis, SIRS – Systemic Inflammatory Response Syndrome, SOFA – Score and Sequential Organ Failure Assessment

Introduction

Sepsis is a systemic reaction of the body to invasive microorganisms such as bacteria and fungi. One of the diseases that patients admitted to the intensive care unit (ICU) may have is to be infected with it (1). Sepsis is the second most common cause of death in patients with non-cardiovascular diseases admitted to the intensive care unit. It is one of the top ten reasons for death among all hospitalized patients.

Sepsis is defined as a syndrome of life-threatening organ dysfunction due to a person‘s dysregulated response to infection. Symptoms include fever, increased heart rate (HR), increased respiratory rate, and decreased consciousness (2). Sepsis is a common disease among children and adults. This disease is among the leading causes of morbidity and mortality in critically ill patients and is the most expensive condition by healthcare spending (3). It has already become a significant global health burden due to higher treatment costs and excessive hospital stays (4). Therefore, it is important to detect sepsis as early as possible. Many sepsis cases result in cardiac arrest (CA) with poor outcomes (5). It has been shown that internationally every year in the world, 30 million people suffer from this disease, and 4.2 million of them are children (6). Among these patients, approximately 750,000 people are with severe sepsis per year, and about one-third of them die (7). For this reason, most recent studies have focused on patients with existing sepsis utilizing electronic medical records, laboratory results, and biomedical signals to predict status changes as sepsis progresses to severe sepsis or septic shock, predict and thus prevent fatal injury and death via intensive management, or analyze the mortality of sepsis patients (8).

In past research, sepsis has been classified into three categories in terms of sepsis progression to severe sepsis to septic shock but recently redefined as two categories in terms of progression from

[1]Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran, and [2]Department of Biomedical Engineering, Faculty of Health, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

Address for correspondence: Mohammad Karimi MORIDANI, Dr, Department of Biomedical Engineering, Faculty of Health, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.

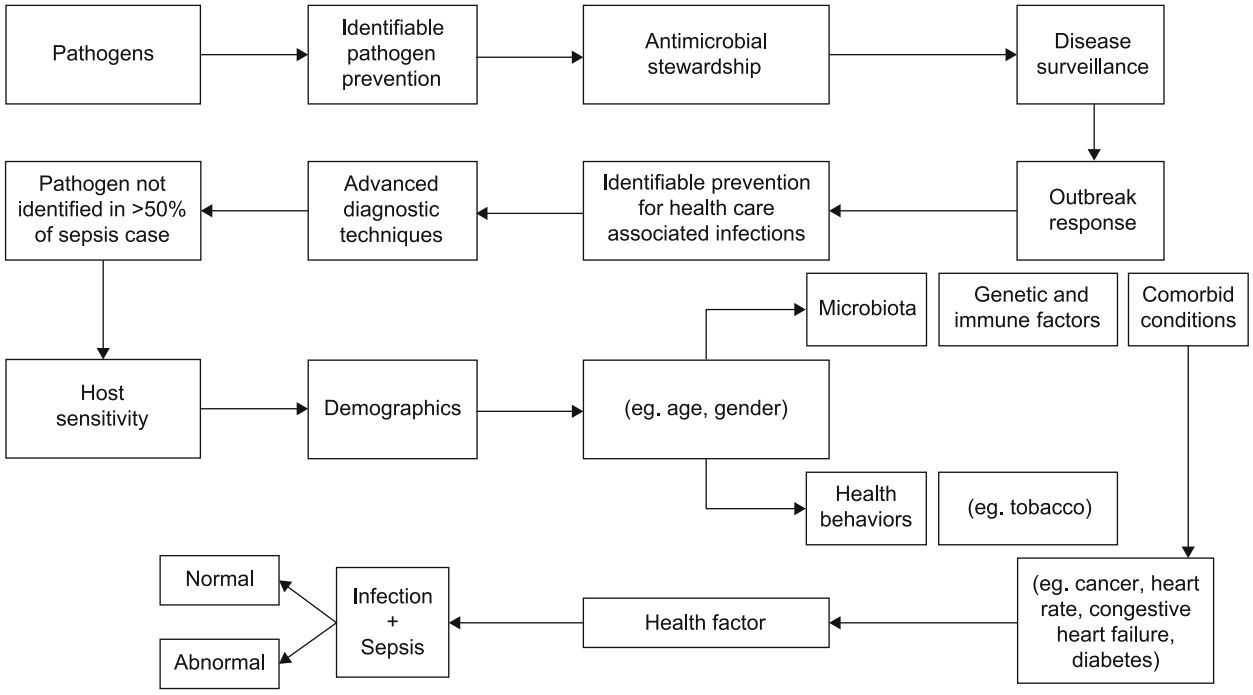

Fig. 1. Block diagram of the prevention of sepsis.

sepsis (encompassing severe sepsis) to septic shock (9). Several studies have shown that early diagnosis and treatment, such as early goal-directed therapy (EGDT), can reduce the risk of severe sepsis and septic shock (10). Systemic inflammatory response syndrome (SIRS) is defined by two or more of the following variables: the temperature is more than 38 ° C and less than 36 ° C, heart rate more than 90 times per minute, respiratory rate returned more than 32 mg per hormone, and or abnormal white blood cell (11). Precise clinical criteria have been reported to identify patients suspected of infection who are at risk of sepsis (12). They identified an episode of suspected infection as the combination of antibiotics and blood cultures within a specific time epoch. They defined the first of these two events as the „onset“ of infection (9). Sepsis accounts for about 25% of intensive care unit (ICU) admissions (13). Patients with sepsis are less likely to archive the return of spontaneous circulation (ROSC) (5). Cardiopulmonary resuscitation in sepsis patients is challenging and unsuccessful. More research is required to prevent CA in these patients (14). The use of continuously measured high-resolution ECG and blood pressure data has provided promising results in the hunt for an accurate predictor. Sepsis is known as a dysregulated immunemediated host response to infection (12).

To identify mortality risk and ensure an appropriate therapeutic interventions, clinical scores have been introduced. In clinical practice, the most commonly used clinical scores are the Acute Physiology and Chronic Health Evaluation (APACHE) II score and the Sequential Organ Failure Assessment (SOFA) score (15). They are validated as the most recognized tools to stratify the severity of the condition. However, with the increasing controversies and complicated methods for using these clinical scoring systems, a growing body of evidence has proposed blood biomarkers as promising alternatives (16). Sepsis diagnostic procedures have been slightly changed since 1991 and include screening labs that may be inaccurate and inaccurate (17). The onset time of sepsis was then defined as an episode of suspected infection with two points or more changes in the Sequential Organ Failure Assessment (SOFA) Score. Using this new definition, Seymour et al. were able to validate the discriminative power of the existing clinical criteria and that transient hypotensive event, identified from the raw blood pressure waveform, which later led to sepsis and higher mortality, were missed by clinical teams for 4 hours on average (18). The key to this discovery was signal quality metrics to reprocess the blood pressure waveform and remove untrustworthy data. Changes in blood pressure and heart rate dynamics are associated with decompensation in critically ill patients (19). Therefore, as shown in Figure 1, a sepsis prevention framework is needed that recognizes patient risk factors and prevention opportunities before the onset of sepsis and the patient presents to the hospital (20). This infection can most often be caused by bacteria or for reasons such as fungi, viruses, or parasites. The severity of the disease determines the outcome of the disease to some extent. One of the most important measures to prevent bloodstream infections is to wash your hands regularly. Quickly clean all wounds, even the smallest wounds. Very important factors affect this disease in different people, including age, sex, etc. If this disease is controlled, it will not cause any problems for the sick person. Otherwise, the person

Tab. 1. Prediction of Sepsis from Clinical Data.

Vital Signs		Laboratory values	
HR	Heart rate (beats per minute)	Glucose Lactate	Serum glucose (mg/dL) Lactic acid (mg/dL)
O2Sat	O2Sat	Magnesium	(mmol/dL)
Temp	Temperature (Deg C)	Phosphate	(mg/dL)
SBP	Systolic BP (mm Hg)	Potassium	(mmol/L)
MAP	Mean arterial pressure (mmHg)	Bilirubin total	Total bilirubin (mg/dL)
DBP	Diastolic BP (mm Hg)	TroponinI	Troponin I (ng/mL)
Resp	Respiration rate (breaths per minute)	Hct	Hematocrit (%)
EtCO2	End tidal carbon dioxide (mmHg)	Hgb	Hemoglobin (g/dL)
Laboratory values		PTT	partial thromboplastin time (seconds)
Base Excess	Measure of excess bicarbonate (mmol/L)	Leukocyte count (count*10^3/µL)	Leukocyte count (count*10^3/µL)
FiO2	Fraction of inspired oxygen (%)	(mg/dL)	(mg/dL)
Ph	N/A	(count*10^3/µL)	(count*10^3/µL)
PaCO2	Partial pressure of carbon dioxide from arterial blood (mm Hg)	Demographics	
SaO2	Oxygen saturation from arterial blood (%)	Age	Years (100 for patients 90 or above)
AST	Aspartate transaminase (IU/L)	Gender	Female (0) or Male (1)
BUN	Blood urea nitrogen (mg/dL)	Unit1	Administrative identifier for ICU unit (MICU)
Alkalinephos	Alkaline phosphatase (IU/L)	Unit2	Administrative identifier for ICU unit (SICU)
Calcium	(mg/dL)	HostAdmTime	Hours between hospital admit and ICU admit
Chloride	(mmol/L)	ICULOS	ICU length-of-stay (hours since ICU admit)
Creatinine	(mg/dL)	Sepsis Label	
Bilirubin direct	Bilirubin direct (mg/dL)	For sepsis patients, Sepsis Label is 1 if t≥tsepsis−6 and 0 if t<tsepsis−6. For non-sepsis patients, Sepsis Label is 0.	

may suffer from critical conditions or even death. In this study, we selected physicians with the help of artificial intelligence to help physicians better and faster diagnose the disease early.

Related works

It is very important to break the diagnosis and treat the disease early if sepsis is diagnosed and treated. This disease is one of the most deadly diseases. In this study, we find the least features and the most important operational features that can handle the data, which is done with confidence and accuracy compared to the work done in this area. Be. The advantage of this is that it makes the network more complex for training and not confidential.

In particular, Mayaud et al demonstrated that heart rate entropy is associated with sepsis in adult critical care subjects (21). Samaneh Layeqian et al are as follows CA-related tasks using machine learning (22). Another paper written by Mohd Basri Mat-Nor is about obtaining a 30-day prediction of sepsis disease by using multiple indicators of function and comparing its performance with the assessment of the failure of successive organs through the SOFA scoring system used (23). Some study presents an algorithm to assess the risk of death in patients with sepsis. In this paper, they used the Simplified Acute Physiology Score (SAPS) for ICU patients and the Sequential Organ Failure Assessment (SOFA) to build their algorithms (24). In the next study written by Lukaszewski et al, the blood of 92 ICU patients were analyzed by RT-PCR

```
HR|O2Sat|Temp|SBP|MAP|DBP|Resp|EtCO2|BaseExcess|HCO3|FiO2|pH|PaCO2|SaO2|AST|BUN|Alkalinephos|Calcium|Chloride|
NaN|NaN|35.1|11.7|22.9|11.9|NaN|311|43.49|0|1|0|-73.68|5|0101.5|96|NaN|131|162.83|NaN|23.5|NaN|NaN|NaN|NaN|NaN
NaN|114|76|NaN|21.5|NaN|NaN|NaN|0.5|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|Na
NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|43.49|0|1|0|-73.68|18|093.5|95.5|NaN|113|76.33|NaN
43.49|0|1|0|-73.68|24|089|98|NaN|106|77|60|19|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN
NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|43.49|0|1|0|-73.68|31|095|97|N
N|NaN|43.49|0|1|0|-73.68|37|096|98|NaN|89.5|70.67|66|23.5|NaN|NaN|NaN|0.5|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|NaN|
```

Fig. 2. An example of a patient's database used in this paper.

HR	O2Sat	Temp	SBP	MAP	DBP	Resp	EtCO2	BaseExcess	HCO3	FiO2	pH	PaCO2	SaO2	AST	BUN	Alkalinephos	Calcium	Chloride	Creatinine	Bilirubin_direct	Glucose	Lactate	Magnesium	Phosphate	Potassium	Bilirubin_total	TroponinI	Hct	Hgb	PTT	WBC	Fibrinogen	Platelets	Age	Gender	Unit1	Unit2	HospAdmTime	ICULOS	SepsisLabel
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	1	1
83	100	NaN	104	62	42	12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	2	1
75	100	35.1	113	71	49	12	NaN	NaN	NaN	0.5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	3	1
70	100	NaN	121	77	50	12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	4	1
70	100	35.1	116	72	45	12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	98	NaN	NaN	NaN	NaN	4.7	NaN	NaN	NaN	29.6	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	5	1
69	100	NaN	98	67	36	12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	98	NaN	NaN	NaN	NaN	NaN	NaN	4.7	NaN	NaN	NaN	29.6	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	6	1
66	100	35.8	111	66	41	12	NaN	NaN	NaN	0.5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	83.95	1	0	1	-2.74	7	1
67	100	NaN	114	68	43	12	NaN	0	20	NaN	7.4	46	99	NaN	69	NaN	8.7	96	3.2	115	NaN	2.3	4.3	4.4	NaN	NaN	NaN	34.8	10.6	NaN	17.5	NaN	289	83.95	1	0	1	-2.74	8	1
68	100	NaN	105	61	39	12	NaN	0	20	NaN	7.4	NaN	99	NaN	69	NaN	NaN	96	NaN	NaN	NaN	2.3	NaN	4.4	NaN	NaN	NaN	34.8	10.6	NaN	17.5	NaN	NaN	83.95	1	0	1	-2.74	9	1

Fig. 3. Display of the sorted data for preprocessing.

expression and neural network analysis of related genes to predict sepsis onset. This study predicted 83.09% of patient cases 1 to 4 days before clinical diagnosis (sensitivity, 91.43 %; specificity, 80.20 %; and accuracy, 94.55 %) (25). Similarly, a study by Jones et al detected the occurrence of sepsis 2 to 3 days before diagnosis by analyzing cell motion using a microfluidic device (26). These methods do not seem to be appropriate because the studies in the previous sentences should be performed daily. Kim et al performed prediction models using a support vector machine (SVM) with temporal features extracted from patient information, such as laboratory tests, biosignal data, and SIRS scores 0-24 h before sepsis diagnosis in 1,239 postoperative patients; 26 patients (2.1 %) had sepsis, and the AUCs ranged between 0.28 and 0.95 (27). In particular, Mayaud et al demonstrated that heart rate entropy is associated with sepsis in adult critical care subjects (21).

Given the high mortality rate associated with sepsis, this article intends to compare the effect of different markers and their relationship with inexpensive and straightforward tests used to diagnose sepsis patients' follow-up in intensive care units.

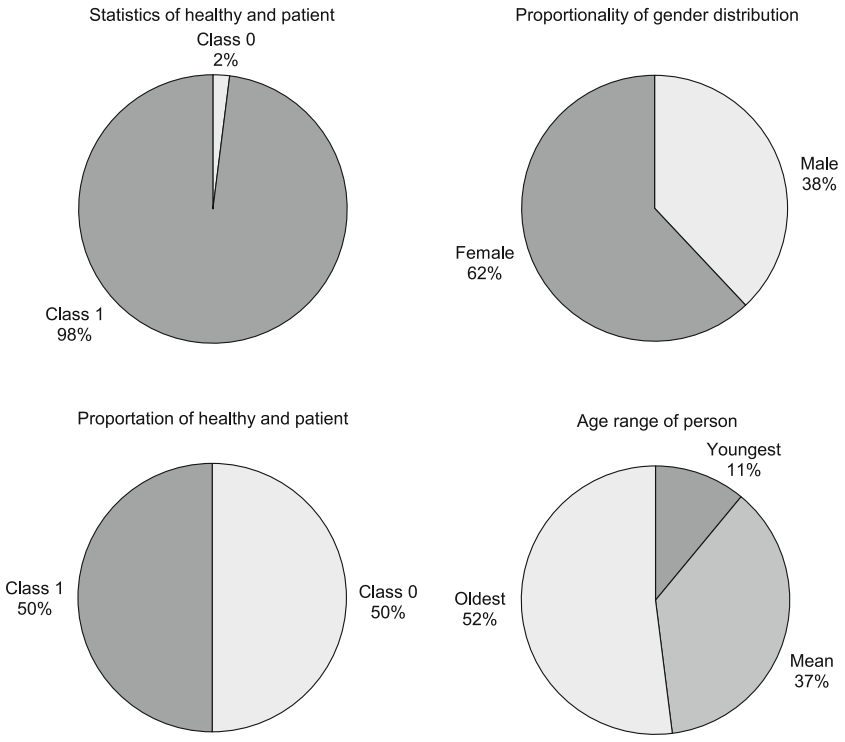

Fig. 4. The statistical description of data.

To continue, this article has been organized as follows:

In the second section, the database and the proposed method to predict sepsis early from clinical data and evaluation methods are discussed. In the third section, the results of the method presented in this article are shown, and also the results of these methods are compared with the features used in this article. Discussion and conclusion are presented in the fourth section.

Materials and methods

A. Describe of dataset

The physio net database is a good resource for researching medical data. The main and continuous mission of this database was to catalyze biomedical research and education. Physio Net also hosts a series of challenges that focus their research on unsolved problems in the clinical and basic sciences. This database contains a collection of discrete, continuous, and medical images. The data we use in this study are discrete.

The goal of this study is the early prediction of sepsis using physiological data. In our clinical data, we have about 41 features, including vital signs, laboratory values, demographics, and sepsis labels, that are shown in Table 1.

In this article, different markers in the diagnosis of sepsis and the effect of each in identifying the disease were studied. Initially, we selected some data out of 40336 data; 50 % of them were healthy, and the rest were patients.

B. Preprocessing

Our data is from ICU patients in a different hospitals. The total number of data is 40336. Sepsis is a time-dependent syndrome that occurs after hours, not days or months. As you can see in Figure 2, these data are taken hourly from the patients whose time intervals were different from each patient, and it averages about 10 hours of data. For healthy people, this period is longer than for people with sepsis. On average, patients were hospitalized for 9 hours and non-patients for 33 hours. For example, Figure 2 shows an example of a patient's data. As shown in Figure 5, the data format is PSV, and we used Notepad++ version 7.6.6.0 and Excel software to sort them. Figure 3 shows the sorted data using this method.

C. Feature selection methods

As mentioned, we have 40336 data indicating people with sepsis and some who do not. Taking the dataset as a whole, the distributions are shown in the upper left of Figure 2, where 1.8 % of all data points have a sepsis label = 1.

In the upper right of Figure 2, the abnormal distribution of data is displayed (patient data accounted for only 0.5 % of total data); we had to select fewer data. We used 50–50 (50 % with sepsis label = 1 & 50% with sepsis label = 0).

The two bottom diagrams in Figure 4 show the data. On average, we have 62 % females and 38 % males. That the youngest is 18 years old, the oldest is 88 years, and their average is 63 years.

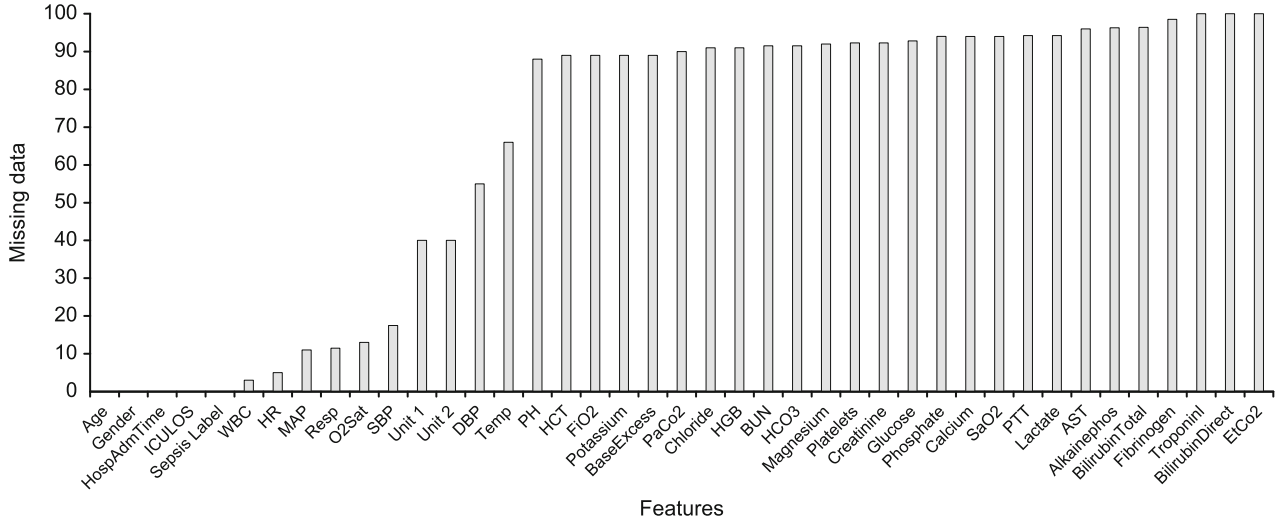

Fig. 5. The Percentage of missing values for each feature.

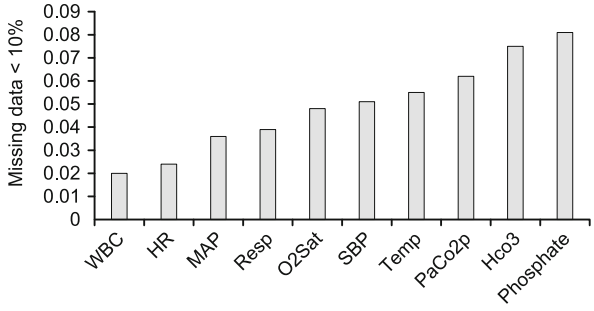

Fig. 6. Graph of the participation rate of each feature.

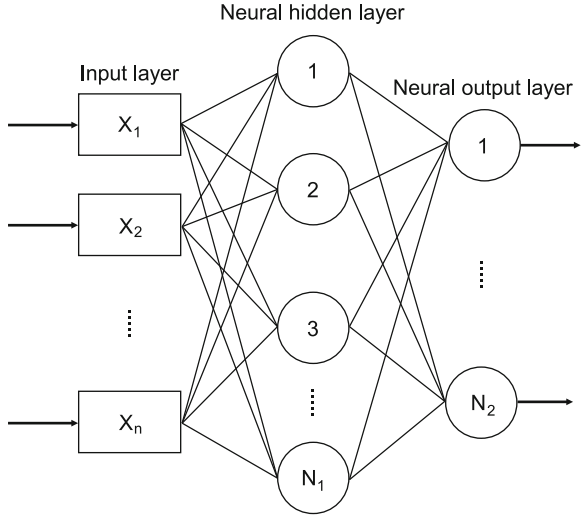

Fig. 7. Structure of MLP Neural Network.

As we can see in Figure 5, we have a lot of missing data that is displayed with Not a Number (NaN). Figure 5 shows the percentage of NaN for each feature. Some features, such as Fibrinogen and hematocrit (HCT), have many missing data and can make modeling more challenging. For this reason, we have used features that have appropriate values, which are mentioned in the method section.

The maximum value of the HO parameter between a healthy person and a patient for a healthy person is 48, and this characteristic has a minimum value of 11 in the patient. The maximum heart rate for a healthy person is 155, and the minimum for a sick person is 41. The minimum temperature recorded for the patient is 32.5 degrees, and the maximum temperature for the patient is 39.25 degrees. Another feature we use is the number of white blood cells in the blood, the maximum of which is 168,200 in sick people and the minimum in patients with 1,300.

We have examined nearly 30 % of the data from all the features listed in Figure 5, with the contribution of each feature by deleting the lost data, as shown in Figure 8. The main target population of our study is shown in Figure 6 that the number of these missing data is below 10 %.

Proposed methods

In this paper, we tried to distinguish healthy subjects from patient data. Important features such as heart rate create meaningful differentiation in some features, but some features need other features to be compared. In this study, behaviors were examined in both patient and healthy groups, and we found that both characteristics are important and effective according to the statistical analyzes performed.

So we used a multilayer perceptron (MLP) to classify healthy and patient. Figure 7 shows a multilayer perceptron consisting of at least three layers of input, output, and processor, also called the hidden layer (28). Except for the input layer segments, each segment consists of a neuron that passes through nonlinear activation

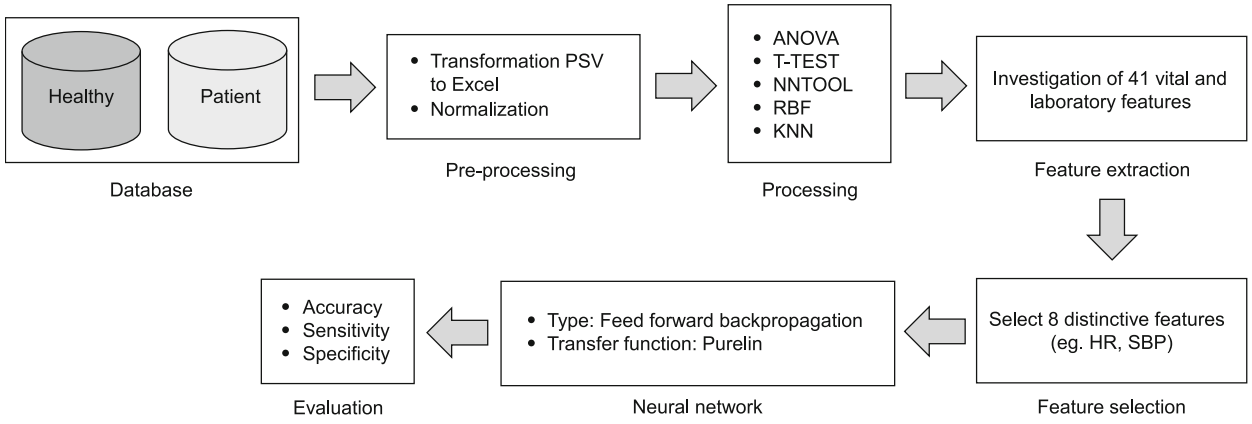

Fig. 8. The block diagram for early prediction of sepsis with clinical data.

functions. MLP uses supervised training called backpropagation for training (28). The artificial intelligence network model performs processing via neurons (29). The presence of multilayers and nonlinear activation functions distinguishes MLP from a linear perceptron. This feature can separate data that is not separated (30). Neural networks can adjust the input parameters if they do not show the optimal response to obtain the desired output. There are several models of learning algorithms to find the relationship between input and output (29).

The multilayer perceptron is also called vanilla when it has a hidden layer (31). Artificial intelligence is used in various areas, such as optimization, modeling, and medical applications (approximation, signal processing, and imaging) (29). The neural network can understand the nonlinear relationship between input and output and cover regression and prediction problems in other fields (32). Artificial intelligence has become more popular in the last two decades because of its high accuracy and speed (29).

Statistical analyzes

According to Figure 5, the database contains intensive care unit data. We need to preprocess the data before processing it. The data format is the pipe-separated value (PSV) and cannot be read in MATLAB software. We converted the PSV format into excel and then normalized it for a better response. Then, we used analysis of variance (ANOVA), t-test, and nntool methods in the processing step. ANOVA is a method of analysis, and a t-test is used to discriminate data. We have 41 data related to vital signs and laboratory properties; then, we only selected 10 of them (feature selection). As indicated in Table 4, among 10 features, 8 of them were differentiated by using the t-test method. Then we used the same 8 features to train and test the network. The network has two layers where the first layer is the input, and the next layer is the hidden layer, which is the first layer consisting of 10 neurons, and the second layer consists of five neurons. Also, the network type is feed-forward backpropagation, and the transfer function is purlin. Then use the criteria of sensitivity, accuracy, and specificity according to formulas (1, 2, 3) to ensure the results of the neural network and to evaluate the effectiveness of the network for the prediction of sepsis. Sensitivity refers to the test's ability to correctly detect ill patients who do have the condition (33). Specificity relates to the test's ability to reject healthy patients without a condition correctly. In the measurement of a set, accuracy refers to the measurements' closeness to a specific value. A true positive (TP) is an outcome where the model correctly predicts the positive class.

Similarly, a true negative (TN) is an outcome where the model correctly predicts the negative class. A false positive (FP) is an outcome where the model incorrectly predicts the positive class. And a false negative (FN) is an outcome where the model incorrectly predicts the negative class. According to formulas (4), in the statistical analysis of binary classification, the F_1 score (also Fscore or F-measure) measures a test's accuracy. Figure 8 shows all the steps used by MATLAB software and other statistical analysis.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Accuracy} = \frac{TN+TP}{TN+FP+TP+FN} \quad (3)$$

In this study, due to being aware of a large number of lost data, we tried to use features that have less than 10 % of the lost data. There was also a small amount of lost data that we preferred to delete.

In block diagram No. 7, we studied the characteristics of many individuals in the extraction block, which recorded about 41 features for each patient, which are described in Table 1. Other parts of the diagram block are stated in the article. These features include heart rate, calcium, age, and other factors, some of which have a lot of missing data. Some features are very important in our studies, such as heart rate and body temperature.

A correlation coefficient is a statistical tool for determining the type and degree of relationship of one quantitative variable. A correlation coefficient is one of the criteria used to determine the

correlation between two variables. The correlation coefficient indicates the severity of the relationship and the type of relationship (direct or inverse). This coefficient is between1 and -1 and is zero if there is no relationship between the two variables. The correlation between two random variables X and Y is defined as follows:

$$\rho_{x,y} = \text{corr}(x,y) = \frac{cov(x\ y)}{\sigma x\ \sigma y} = \frac{E[(x-\eta x)(y-\eta y)]}{\sigma x \sigma y} \quad (4)$$

E is the mathematical expectation operator, cov means the covariance and corr is the usual symbol for Pearson's correlation, and sigma is the standard deviation symbol.

Based on the correlation coefficient analysis described in Table 2, the criterion for values that have greater independence, lower dependence, and less effectiveness than each other is a value of less than 0. 5. for example, the correlation coefficient for the two WBC properties and (Correlation < 0.05) Temp is considered the criterion for investigating other parameters. Properties whose correlation coefficients are closer to zero have more independent conditions than those that tend toward the number. Another example is comparing the two HR and MAP properties, which is not a good feature compared to other parameters because this property's correlation coefficient tends to be a number. A comparison of SBP and WBC is more accurate because its correlation coefficient tends to zero.

In the science of impact size statistics, a so-called quantitative measure of the magnitude of a phenomenon. Examples of effect size are the correlation between two variables, the regression coefficient in a regression, and the mean of the difference or even the hazard that occurs. Like some people die of sepsis, and some recover. For most effect sizes, the larger absolute value always indicates a stronger effect. Formula 5 shows how to calculate the d effect size.

The correlation coefficient is a statistical tool to determine the type and degree of relationship between one quantitative variable and another quantitative variable. The correlation coefficient is one of the criteria used to determine the correlation of two variables. The correlation coefficient indicates the intensity of the relationship as well as the type of relationship (direct or inverse). This coefficient is between 1 and -1, and if there is no relationship between the two variables, it is equal to zero.

$$d = \frac{m1 - m2}{\sqrt{\frac{\sigma 1^2 + \sigma 2^2}{2}}} \quad (5)$$

Where *m* is the mean of the study group, and σ represents the studied groups' variance. Table 3 shows the effect size results of the extracted parameters between the two groups.

The effect size index is independent of the sample size and shows the relationship between the two variables. This test is used to determine the significance of the difference between the independent samples. The effect size reduces the size of the difference between the two groups. One of the methods for recognizing the effect size is Cohen d method, which shows the difference between the two divisors by the standard deviation of the data. The larger the effect size, the greater the difference between the groups. For example, the relationship between SBP and HR, which has a value of 1.41 according to Table 3, is stronger and more distinct than the relationship between Resp and H_3O_3, which has a value of 0.100.

Tab. 2. Selected heat map package. The heat map was generated based on 10 features from a normal dataset.

	HR	O_2Sat	SBP	H_3O_3	Phosphate	WBC	MAP	Resp	$PaCo_2$	Temp
HR	1	-0.33	-0.24	0.24	-0.14	-0.14	0.96	0.06	0.23	-0.05
O_2Sat	-0.33	1	0.14	0.26	0.13	0.31	-0.30	-0.33	0.16	-0.004
SBP	-0.24	0.14	1	-0.14	-0.009	0.07	-0.27	-0.01	-0.14	-0.12
H_3O_3	0.24	0.26	-0.14	1	0.14	0.17	0.23	-0.10	0.52	-0.08
Phosphate	-0.14	0.13	-0.009	0.14	1	0.11	-0.15	-0.20	-0.02	0.04
WBC	-014	0.31	0.07	0.17	0.11	1	-014	-0.04	-0.12	0.05
MAP	0.96	-0.30	-0.27	0.23	-0.15	-0.14	1	0.02	0.24	-0.04
Resp	0.06	-0.33	-0.01	-0.10	-0.20	-0.04	0.02	1	0.01	0.26
$PaCo_2$	0.23	0.16	-0.14	0.52	-0.02	-0.12	0.24	0.01	1	0.09
Temp	-0.05	-0.004	-0.12	-0.08	0.04	0.05	-0.04	0.26	0.09	1

Tab. 3. Physiological parameters analyzing by Cohen's d effect size between two groups.

	HR	O2Sat	SBP	H3O3	Phosphate	WBC	MAP	Resp	PaCO2	Temp
HR	0	2.53	1.41	11.07	15.29	13.52	2.36	10.85	4.79	10.73
O2Sat	2.53	0	3.00	14.97	33.40	22.69	1.00	14.29	3.90	22.82
SBP	1.41	3.00	0	8.11	9.91	9.22	3.03	8.03	4.94	7.39
H3O3	11.07	14.97	8.11	0	4.52	2.46	5.45	0.10	1.84	3.51
Phosphate	15.29	33.40	9.91	4.52	0	2.38	7.66	4.44	3.51	28.55
WBC	13.52	22.69	9.22	2.46	2.38	0	6.80	2.49	2.87	9.61
MAP	2.36	1.00	3.03	5.45	7.66	6.82	0	5.36	2.34	4.47
Resp	10.85	14.29	8.03	0.100	4.44	2.49	5.36	0	1.78	3.19
PaCo2	4.79	3.90	4.94	1.84	3.51	2.87	2.34	1.78	0	0.75
Temp	10.73	22.82	7.39	3.51	28.55	9.61	4.47	3.19	0.75	0

Neural network algorithms

This section optimized the features with statistical analysis that was processed through different perceptron neural networks, radial basis function (RBF), and K- Nearest Neighbor (KNN). The RBF network is very convenient for the interpolation network method, where the k-means algorithm is to be used for clustering. When training networks. Anyone of the test matrix examples closest to the cluster's center in the learning matrix is classified as that cluster's class.

$$y(x) = \sum_{i=1}^{N} w_i \phi(||x - x_i||) \quad (6)$$

Formula 6, grid input x parameter, parameter ϕ is the Gaussian activation function, N is the number of hidden layer neurons, x_i is the center vector of neuron i, and w is the weight of neuron i in the linear output neuron.

$$D_{euc} = \sum_{i=1}^{p} (x_i - y_i)^2)\text{^}1/2 \quad (7)$$

Formula 7 represents the two variables x and y, including two points in Euclidean space and p, which also represent the space in question. The final result of the Euclidean relationship is never negative.

Unlike the other networks mentioned, this network does not require training, and the classes in this network are single-columned. The classification method works because each example of the test matrix will be compared with an example of a trained matrix, and any one that is the closest will appear in that class. Our criterion for testing in this method is the Euclidean formula. According to Table 1, with the nearest neighbor number's change, we could get different outcomes, with the best nearest neighbor number being 1. The multilayer perceptron is a part of neural network feedback. An MLP (multilayer perceptron) consists of three layers of nodes: an input layer, a hidden layer, and an output layer. Other than the input nodes, additional nodes are each a neuron that uses a nonlinear activation function. If an MLP (multilayer perceptron) holds a linear activation function in each neuron, technically, weighted inputs are drawn with this same linear function. The activation function used in this experiment was the hyperbolic tangent function represented as follows:

$$Y(v_i) = \text{atanhb (bv)} \quad (8)$$

Parameter V is the weighted sum that is entered in the activation function and helps to calculate the output.

In this function, a = 1.7951 and b = 2 / 2.

Learning in the neural network occurs with the change of connection weights after processing each piece of data, based on the output error rate compared with the expected result from beforehand. The error in the output node j in the Nth point is denoted as $t_j(n) = d_j(n) - y_j(n)$ where t is the target value and y is the value produced by the perceptron.

$$\varepsilon(n) = \frac{1}{2}\sum_{j} e_j^2(n) \quad (9)$$

By using the gradient, the variation in weight is as follows:

$$-\frac{\delta\varepsilon(n)}{\delta_{vj}(n)} = e_j(n)\phi'(v_j(n)) \quad (10)$$

Where, ϕ' is the derivative of the activation function.

Results

According to our study, different markers in the diagnosis of sepsis and the effect of each in identifying the disease were examined. We used features such as mean, standard deviation (std), variance (var), median, mode, skewness, and kurtosis in MATLAB software with a neural network. Unfortunately, overall, the neural network responses were not appropriate because using a lot of features, the complexity of network computing increased, and ifferentiation between the two groups decreased.

In the next step, we selected the optimum feature and only used the mean. Finally, according to Table 4 to identify optimal data, we used a T-test with a p-value less than 0.05 ($p < 0.05$), DE Kohen effect size. It also shows the three features of specificity, sensitivity, and accuracy. We used Boxplot in MATLAB software to show more distinction, as shown in Figures 9 and 10.

As the study in this paper shows, the four statistical features studied at each stage, such as variance, median, mean, and std, are more distinct. For example, we found a significant difference in the heart rate variance for the two healthy and patient populations, which is a numerical difference of 45.84. In the next feature, when we study the o2sat of healthy and patient people, we find out that there is a difference in fashion, which equals 5. According to the study, the next feature of SBP in healthy individuals compared to patients who have a median difference of 34.5. The last differentiating feature that was examined was the body temperature characteristic of healthy and healthy subjects, which, as evidenced by the std, had a difference of 1.26.

This section examines healthy and patient groups in traits that do not differ and will not help our study. As can be seen, the two

Tab. 4. The statistical method using t-test analysis.

Clinical Variables	Baseline	Septic	p
HR	112.84	82.27	7.0401×10^{-36}
Temp	37.13	36.38	7.5490×10^{-5}
O2Sat	93.87	96.98	5.6791×10^{-10}
SBP	133.85	108.07	1.4817×10^{-11}
HCO3	22.58	23.14	0.4662
Phosphate	3.50	4.31	0.0015
WBC	10.52	16.61	5.3372×10^{-4}
MAP	83.23	70.24	5.1441×10^{-9}
Resp	23.18	20.88	0.0136
PaCO2	46.00	45.90	0.9656

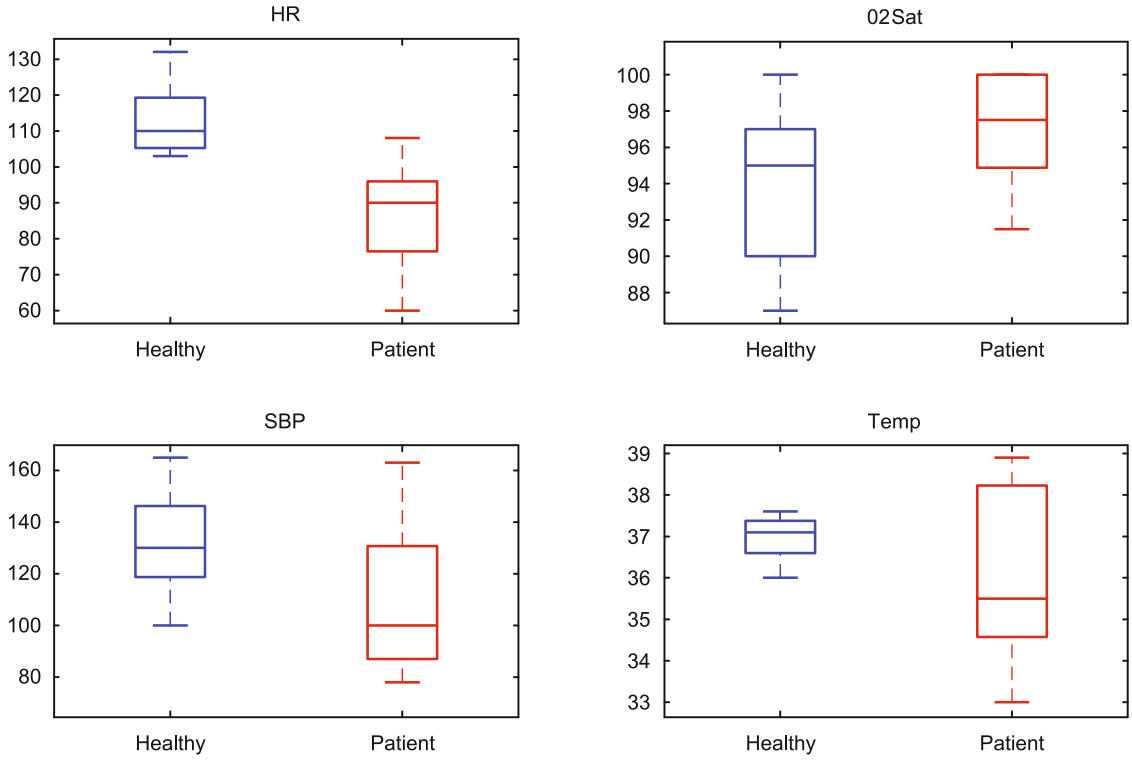

Fig. 9. Showing a difference between two groups using a boxplot.

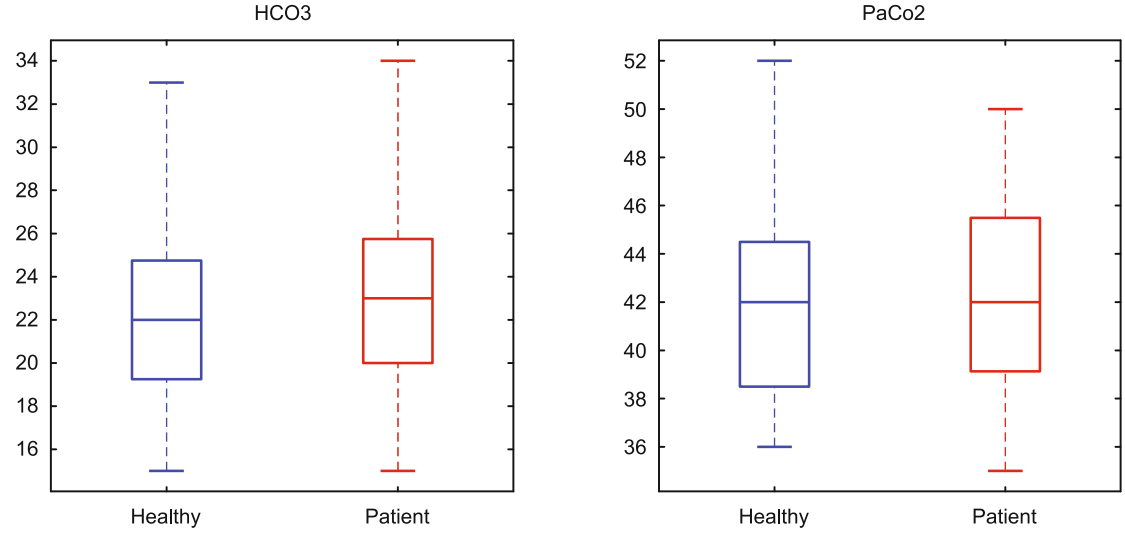

Fig. 10. Showing a difference between two groups by a boxplot.

critical properties of HCO_3 and $PaCO_2$ are not very different in this study. In all statistical properties, we can say that they are almost equivalent and have no significant difference.

Table 5 is related to Figure 9 and Figure 10, showing the differentiation of healthy and patient samples with var, median, std, and mod properties. For example, in Figure 2, the variance of the HR for healthy people is 99.68, and patient variance is 145.52, and in Figure 2, $SatO_2$ for healthy people is 95, and patient mode is 100, and for Figure 8, SBP and Temp are shown in Table 5. But as it is evident in Figure 7, there is little differentiation in HCO_3 and $PaCO_2$, and healthy and patient mean and variance are close to each other, so it is not a suitable parameter for our analysis. So criterion evaluation is as follows in Table 5 with the mean and standard deviation 5 ± 3.

Tab. 5. Comparison between two groups based on Statistical features.

Feature	Healthy				Patient				Healthy		Patient	
	HR	O2Sat	SBP	Temp	HR	O2Sat	SBP	Temp	H3CO3	PaCO2	H3CO3	PaCO2
var	99.68	12.16	342.95	0.27	145.52	13.88	683.92	3.19	33.13	289.09	26.62	249.87
median	110	95	135.50	36.95	86	98	101	35.80	22	42	23	42
std	9.98	3.48	18.51	0.52	12.06	3.72	26.14	1.78	5.75	17	5.16	15.80
mode	103	95	122	36.67	88	100	94	35.83	22	43	23	32
p	<0.05								>0.05			

Tab. 6. Evaluation criteria.

Network type	NumNeighbour	Cluster Size	Hidden Layer 1	Hidden Layer 2	Hidden Layer 3	Sensivity	Specifity	Accuracy
KNN	2	–	–	–	–	95.19%	100%	97.59%
KNN	3	–	–	–	–	98.07%	100%	99.03%
KNN	4	–	–	–	–	94.23%	100%	97.11%
KNN	5	–	–	–	–	94.23%	99.03%	96.63%
KNN	6	–	–	–	–	91.34%	98.07%	94.71%
RBF	–	8	–	–		96.77%	100%	98.41%
MLP	–	–	10	–	–	90.62%	100%	95.31%
MLP	–	–	7	5	–	96.87%	100%	98.43%
MLP	–	–	7	6	2	87.5%	100%	93.75%

Tab. 7. Comparison of results.

	Sensitivity	Specificity	Accuracy
SOFA	70%	59%	64.5%
Q-SOFA	92%	85%	88.5%
KNN	98%	100%	99%
RBF	96%	100%	98%
MLP	96%	100%	98%

Finally, concerning the attributes given to the network and described in the method section, the results are shown in Table 6, showing that the KNN network with neighborhood number 1 and accuracy 1.2, in the RBF network with Cluster Center 2 and Accuracy 2.9, and in two-layer perceptron, which includes the input layer with two neurons and the first hidden layer with two neurons with an accuracy of 0.8, we achieved the best results among which the KNN network performs best. Table 6 details the neural network results.

Table 7 explains the results of the neural network in detail. In Table 6, the learning rate for the perceptron neural network with different layers is set to 0.0001 for all modes.

Sepsis is one of the most common causes of death among patients in the intensive care unit worldwide. Despite new supportive therapies and strong antibiotics, sepsis is still a risk factor in patients‘ lives. This paper aimed to compare the value of some physiological parameters in the diagnosis of sepsis and investigate their relationship to select a more practical diagnostic method.

In this study, statistical analysis of P-value, Anova, and De-Cohen effect size is used, which is a low-cost and easy way to find suitable features for research. Using these statistical methods is to find the best feature for processing in the neural network. The data in this study consisted of 41 attributes that were analyzed by eight features. After identifying the appropriate features using the mentioned analyses to detect the disease, the features were assigned to single-layer, double-layer, and three-layer perceptron neural networks, the best response with different epochs equals to the lowest error. High precision is shown to us at the output. In addition to the network with different layers, we investigated the RBF network features with a center of gravity smaller than the number of features and the Euclidean activation function. M.A Baig et al in an article tried to determine the mortality rate of sepsis from the SOFA and Q - SOFA difference, comparison to our paper‘s results is in Table 7.

Discussion

As shown in Table 8, we have reviewed previous research on sepsis infection detection using neural networks and compared our results with them. In 2020, Jonathan Freund et al redefined the concept of sepsis with an international working group. To identify patients at risk for mortality, the task force recommended rapid organ failure assessment scores instead of systemic inflammatory response syndrome criteria. Out of 1088 patients screened, 879 patients were analyzed. The mean medical age was 67 years, ranging from 47 to 81 years, 414 (47 %) were female, and 379 (43 %) had a respiratory infection. The in-hospital mortality rate was 8 %: 3 % for patients with a quick sepsis-related organ failure assessment (qSOFA) score lower than two versus 24 % for those with a qSOFA score of 2 or higher. qSOFA performed better than SIRS and severe sepsis in predicting in-hospital mortality (34). In this study, a systematic review and meta-analysis were performed to assess prospective integration accuracy in patients with suspected sepsis. A comprehensive electronic search was conducted through the Internet retrieval system as of December 15, 2014. Methodological quality assessment was performed using the QUADAS2 tool.

Tab. 8. Comparison of the results of different methods in previous studies.

Author	Year	Method	Results		
Jonathan Freund et al (34)	2020	QSOFA	The area under the receiver performance curve = 80 %		
Jiawan Wu et al (35)	2015	QUADAS2	sensitivity	specificity	SORC
			78 %	83 %	89 %
Zhong Zheng et al (36)	2015	Meta-analysis method	sensitivity	specificity	SORC
			77%	73 %	85 %
Fabian Jaimes (15)	2005	Logistic regression and neural network	SORC for NN		SORC for LR
			87 %		75 %
Jau-Woei Perng et al (37)	2019	Deep Learning Stimulated neural network plus SoftMax	Accuracy (28 days)		
			81.59 %		
Yuan Luo (38)	2017	LSTM (Long Short-Term Memory)	precision	recall	f-measure
			72 %	68 %	70 %
Roberts et al (39)	2011	SVM (support vector machine)	precision	recall	f-measure
			72 %	75.3 %	73.7 %

The diagnostic value of perspective in sepsis was evaluated using a mixture of sensitivity, specificity, probability ratio, odds ratio, and a summary of the receptor performance characteristics curve. The susceptibility of perspective to sepsis was 0.78. The mixture specificity was 0.83, the positive probability ratio was 4.63, the negative probability ratio was 0.22, the mixed odds ratio was 21.73, the area under the receptor function summary curve was 0.89, and the Q index was 0.82. This meta-analysis shows that perspective has a special advantage in inpatient management and may be a useful and valuable marker in early sepsis diagnosis. However, perspective showed moderate diagnostic accuracy in distinguishing sepsis from non-sepsis, which precluded its recommendation as a final test for sepsis diagnosis in isolation (35). Zhong Zhen et al in 2015 in a study aimed to systematically and quantitatively evaluate the value of perspective for the diagnosis of sepsis using meta-analysis. A total of eight studies, including 1757 patients, were included in this meta-analysis. Sensitivity, specificity, and diagnostic odds ratios were 0.77, 0.73, and 14.25, respectively. The characteristic curve area of the receiver operating factor (SROC) below the curve was 0.8585. Subgroup analysis excluding deprivation of outdoor environments showed that sensitivity and specificity were 0.85 and 0.65, respectively. Perspective in combination with other laboratory biomarkers in the diagnosis of sepsis may focus on future studies (36). Neural networks are a new methodological tool based on nonlinear models. They appear to be better at predicting and classifying biological systems than traditional strategies such as logistic regression. This article provides a practical example that contrasts with both approaches to sepsis‘s suspected presence in the emergency room. The statistical population includes patients suspected of bacterial infection as their primary diagnosis for emergency hospitalization in two hospitals located at the university. A total of 533 patients were selected, and the 28-day mortality was 19 %. The network included all variables, and there was no significant difference in predicting between approaches. The active areas below the characteristic receptor curves for the logistics and neural network models were 0.7517 and 0.8782 (p = 0.037), respectively. A predictive model can be a useful tool for creating suspected sepsis in the room (15).

In the study by Jau-Woei et al, a deep learning algorithm was used to predict the mortality of suspected infected patients in a hospital‘s emergency department. In January 2007 and December 2013, the 4,220 patients included in this study were admitted to the emergency department due to suspected infection. In the present study, an in-depth learning structure was developed to predict mortality in septic patients and compared with several machine learning methods and two sepsis screening tools: SIRS and qSOFA, as a predictor of mortality for septic patients who died within 72 hours and 28 days. The results showed that the accuracy of deep learning methods, significantly stimulated neural network plus SoftMax (87.01 %) in 72 hours and 81.5 9% in 28 days), is higher than other device learning methods, SIRS and qSOFA. We expect in-depth learning to effectively assist medical staff in the early detection of sensitive patients (37).

Conclusion and future work

This study included the diagnosis of sepsis using machine learning. In this study, using the neural network, train the radial base’s function, one-layer, two-layer, and three-layer perceptron, K, the nearest neighbor of the network, and finally test. Due to the features available in this database, we over-budgeted the use of P-Value and T-Test and used the De-Kohen-size criterion to select and rank the standard features used, finally using selected features in the network for training achieved acceptable results that can be seen in the table (your article). Table (6) with several selected network features (6) has the best quality and accuracy.

Due to the pathophysiological complexity of the infectious disease and the involvement of many inflammatory mediators in it, a combination of biomarkers may be used to make the diagnosis, monitoring, and prediction of disease outcomes more effective. The proposed method is much more economical and can help physicians to treat patients. Besides, clinical tests performed in laboratories and hospitals can be omitted. This paper introduces sepsis biomarkers that can help identify patients, evaluate response to treatment, distinguish systemic sepsis from local, and even assist clinicians in differentiating sepsis patients from patients with non-infectious SIRS.

In this study, we have achieved acceptable results by reducing the characteristics to 5 characteristics. Our advice to future researchers who are interested in research in this field is to reduce the features further and use other methods for training and testing in Machine learning can help improve results and help professionals.

References

1. Deborah J, Stearns-Kurosawa MFO. The Pathogenesis of Sepsis 2011; 6: 19–48.

2. Fethi Gül MKA, Cinel I, Kumar A. Changing Definitions of Sepsis. Turk J Anaesthesiol Reanim 2017.

3. Shannon A, Novosad MRPS, Grigg C, Lake J, Robyn M. Epidemiology of Sepsis: Prevalence of Health Care Factors and Opportunities for Prevention. 2016.

4. Gaieski DF et al. Benchmarking the incidence and mortality of severe sepsis in the United States. Crit Care Med 2013; 41 (5): 1167–1174.

5. Morgan RW et al. Sepsis-associated in-hospital cardiac arrest: epidemiology, pathophysiology, and potential therapies. J Crit Care 2017; 40: 128–135.

6. Rudd KE et al. Global, regional, and national sepsis incidence and mortality, analysis for the Global Burden of Disease Study. Lancet 2020; 395 (10219): 200–211.

7. Stevenson EK et al. Two decades of mortality trends among patients with severe sepsis: a comparative meta-analysis. Crit Care Med 2014; 42 (3): 625.

8. Henry KE et al. A targeted real-time early warning score (TREWScore) for septic shock. Science Trans Med 2015; 7 (299): 299ra122–299ra122.

9. Singer M et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 2016; 315 (8): 801–810.

10. Nguyen HB et al. Implementation of a bundle of quality indicators for the early management of severe sepsis and septic shock is associated with decreased mortality. Crit Care Med 2007; 35 (4): 1105–1112.

11. Levy MM et al. sccm/esicm/accp/ats/sis international sepsis definitions conference. Intens Care Med 2003; 29 (4): 530–538.

12. Shashikumar SP et al. Early sepsis detection in critical care patients using multiscale blood pressure and heart rate dynamics. J Electrocardiol 2007; 50 (6): 739–743.

13. Esteban AM, Frutos-Vivar FM. Sepsis incidence and outcome: Contrasting the intensive care unit with the hospital ward. 2007; 5 (35).

14. Javan SL, Sepehri MM. Toward analyzing and synthesizing previous research in early prediction of cardiac arrest using machine learning based on a multi-layered integrative framework. J Biomed Inform 2018; 88: 70–89.

15. Ferreira F, Bota DP. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA 2001; 286: 1754–1758.

16. Vincent SM. Ten reasons why we should NOT use severity scores as entry criteria for clinical trials or in our treatment decisions. Crit Care Med 2010; 38 (1): 283–287.

17. Calvert JS et al. A computational approach to early sepsis detection. Comp Biol Med 2016; 74: 69–73.

18. Hug CW, Clifford GD. Clinician blood pressure documentation of stable intensive care patients: an intelligent archiving agent has a higher association with future hypotension. Crit Care Med 2011; 39 (5): 1006.

19. Li-wei HL, Mark RG. A model-based machine learning approach to probing autonomic regulation from nonstationary vital-sign time series. IEEE J Biomed Health Inform 2016; 22 (1): 56–66.

20. Dantes RB, Epstein L. Combatting sepsis: a public health perspective. Clin Infect Dis 2018; 67 (8): 1300–1302.

21. Mayaud L et al. Predictive power of heart rate complexity to estimate severity in severe sepsis patients. J Crit Care 2013; 6 (28): 37.

22. Javan SL et al. An intelligent warning model for early prediction of cardiac arrest in sepsis patients. Comp Methods Programs Biomed 2019; 178: 47–58.

23. Shukeri et al. Sepsis mortality score for the prediction of mortality in septic patients. J Crit Care 2018; 43: 163–168.

24. Bouadma L. Use of procalcitonin to reduce patients‘ exposure to antibiotics in intensive care units (PRORATA trial): a multicentre randomised controlled trial. 2010; 375 (9713): 463–474.

25. Lukaszewski RA et al. Presymptomatic prediction of sepsis in intensive care unit patients. Clin Vaccine Immunol 2008; 15 (7): 1089–1094.

26. Jones CN et al. Spontaneous neutrophil migration patterns during sepsis after major burns. PloS One 2014; (12): e114509.

27. Kim J, Blum J, Scott C. Temporal features and kernel methods for predicting sepsis in postoperative patients. Technical report, Citeseer. 2010.

28. Rosenblatt F. Principles of neurodynamics. perceptrons and the theory of brain mechanisms, Cornell Aeronautical Lab Inc Buffalo NY. 2016.

29. Ghritlahre HK, Prasad RK. Application of ANN technique to predict the performance of solar collector systems – A review. Renewable Sustainable Energy Rev 2018; 84: 75–88.

30. Cybenko G. Approximation by superpositions of a sigmoidal function. Mathemat Control, Signals Systems 2003; 5 (4): 455–455.

31. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction: Springer Science & Business Media. 2009.

32. Yesilnacara TT. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study. Hendek region (Turkey) 2005; 79: 251–266.

33. Drum DE, Christacopoulos JS. Hepatic scintigraphy in clinical decision making. J Nuclear Med 2004; 13 (12): 908–915.

34. Freund J, Lemachatti N. Prognostic Accuracy of Sepsis-3 Criteria for In-Hospital Mortality Among Patients With Suspected Infection Presenting to the Emergency Department. JAMA 2017; 3: 301–308.

35. Wu JY, Zhang GH, Wu FP.Taiping He Accuracy of Presepsin in Sepsis Diagnosis: A Systematic Review and Meta-Analysis. 2015.

36. Zheng ZJ, Ye LG, Gao YZ, Tang LP, Zhang M. The accuracy of presepsin for the diagnosis of sepsis from SIRS: a systematic review and meta-analysis. 2015.

37. Perng JW, Kung CT. Mortality Prediction of Septic Patients in the Emergency Department Based on Machine Learning. 2019.

38. Luo Y. Recurrent neural networks for classifying relations in clinical notes. J Biomed Inform 2017.

39. Rink B, Roberts K. Automatic extraction of relations between medical concepts in clinical texts. 2011.

Received September 12, 2022.
Accepted October 11, 2022.