

RESEARCH

Open Access



Applying artificial neural network for early detection of sepsis with intentionally preserved highly missing real-world data for simulating clinical situation

Yao-Yi Kuo¹, Shu-Tien Huang² and Hung-Wen Chiu^{3*}

Abstract

Purpose: Some predictive systems using machine learning models have been developed to predict sepsis; however, they were mostly built with a low percent of missing values, which does not correspond with the actual clinical situation. In this study, we developed a machine learning model with a high rate of missing and erroneous data to enable prediction under missing, noisy, and erroneous inputs, as in the actual clinical situation.

Materials and methods: The proposed artificial neural network model was implemented using the MATLAB ANN toolbox, based on stochastic gradient descent. The dataset was collected over the past decade with approval from the appropriate institutional review boards, and the sepsis status was identified and labeled using Sepsis-3 clinical criteria. The imputation method was built by last observation carried forward and mean value, aimed to simulate clinical situation.

Results: The mean area under the receiver operating characteristic (ROC) curve (AUC) of classifying sepsis and nonsepsis patients was 0.82 and 0.786 at 0 h and 40 h prior to onset, respectively. The highest model performance was found for one-hourly data, demonstrating that our ANN model can perform adequately with limited hourly data provided.

Conclusions: Our model has the moderate ability to predict sepsis up to 40 h in advance under simulated clinical situation with real-world data.

Keywords: Sepsis, Prediction, Artificial neural network, Machine learning, Artificial intelligence

Introduction

Sepsis is a clinical syndrome caused by a dysregulated host response to infection [1]. This inflammatory response can lead to multiple organ dysfunction syndrome, including acute respiratory distress syndrome, acute renal failure, disseminated intravascular

coagulation, and even death. For decades, sepsis has been considered challenging to treat in hospitals globally given its high mortality and high medical costs. Older patients aged ≥ 65 years account for the majority (60–85%) of all cases of sepsis, as older people are more susceptible to infection and have a higher risk of sepsis [2–5]. With the older population increasing worldwide, the incidence of sepsis may continue to increase, resulting in sepsis being a persistent, challenging problem. Identifying early sepsis, an early form of infection, is important to prevent sepsis progressing to severe condition such as severe

*Correspondence: hwchiu@tmu.edu.tw

³ Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

sepsis or septic shock; each hour of delayed treatment is associated with an approximately 3.6–9.9% increase in mortality [6]. Furthermore, in areas of the world with the lowest socio-demographic index, the need for greater prevention of sepsis is highlighted by Global Burden of Disease Study 2017 [7], which emphasized the need to identify and to predict sepsis. However, no formal definition exists for early sepsis. Conflicting results have been provided for the ability of warning scores systems such as Quick Sepsis-Related Organ Failure Assessment (qSOFA) and National Early Warning Score (NEWS) to predict early sepsis [8–10]; patients may be misclassified as having sepsis based on their inflammation status, leading to a higher rate of antibiotic use and *Clostridioides difficile* infection, and antibiotic use did not affect 30-day mortality [11]. Moreover, in patients with systemic inflammatory response syndrome (SIRS) without evidence of infection, sepsis could not be predicted or identified, as SIRS is not always caused by infection [12].

Rapid progress has been made in machine learning in the last few years. Machine learning involves computer programs that undergo a learning process, with different rules attempted and learning performance improved. Machine learning is an influential and powerful tool for turning information into knowledge and is good at learning the rules governing a phenomenon [13]. Some studies have applied machine learning for data mining for diagnosing appendicitis [14] and diabetes [15] and for tumor assessment [16].

Applying machine learning with diverse variables and indicators has also been investigated. Akram's team used continuous (minute-by-minute) physiologic data to predict sepsis and demonstrated that salient physiologic markers are temporally and differentially expressed in septic patients [17]. Joseph's team used physiologic, laboratory data and subjective variables to predict onset of vasopressor therapy and found that practice-specific features denoting measurement recency improved local performance [18]. An-Kwok's team discussed the use of an expansive number of physiologic, laboratory, and demographic variables to create efficient, automated prediction of acute respiratory failure and acute respiratory distress syndrome [19].

Some predictive systems using machine learning models have been developed to predict or identify sepsis [20]. Gradient tree boosting models with 0% missing inputs using only vital signs can achieve the performance of 0.90 area under receiver operating characteristic (ROC) curve (AUC) when identifying sepsis and can achieve the performance of 0.84 AUC when predicting sepsis 24 h prior to onset [21, 22]. Logistic regression models using laboratory data with 7% missing inputs can achieve the performance of 0.83 AUC when identifying sepsis [23]. These

models achieved favorable performance in the presence of low-percent missing and erroneous data. Nonetheless, in the actual clinical situation, missing and erroneous data exist due to several reasons. Some studies have reported that these missing and erroneous data have become a challenge for machine learning models to convert information into knowledge. The AUC of a gradient tree boosting model can decrease from 0.90 to 0.75 in the presence of 60% missing data [21].

Artificial neural network (ANN), a machine learning model, has been successfully used to solve highly difficult and complex problems in the field of physical sciences and in organizational research. ANN enables faster and efficient data collection and processing [24, 25]. Furthermore, as it is regarded as a practical and flexible modeling tool, ANN can generalize pattern information to new data, and it has information processing characteristics to learning power, high parallelism, fault tolerance, nonlinearity, noise tolerance, and capabilities of generalization [25]. One study used ANN to classify bacteremia and nonbacteremia patients with 20 clinical variables, including demographic variables, vital signs, and laboratory data. The AUC of prediction performance was 0.729 (95% confidence interval [CI]: 0.712–0.728) [26]. Another study used ANN for neonatal sepsis diagnosis with 25 maternal and neonatal features. The prediction performance was 0.933 in sensitivity, 0.800 in specificity and 0.944 in AUC [27].

Above published models could discriminate between sepsis and nonsepsis patients. However, a reliable model should be established to predict the sepsis onset timing in advance using before-sepsis-onset data with a high missing rate, corresponding to the actual clinical situation. Therefore, in this study, we developed a model based on ANN for sepsis prediction by using patient vital signs and laboratory data comprising up to 80% missing and erroneous data as the input, to see an easy shallow network is suitable to address these problems or not. First, we used ANN to classify sepsis and nonsepsis patients with different sepsis onset timings prior to onset. Second, we assessed how the different timings prior to onset affect prediction performance. Finally, we attempted to precisely predict the timing of sepsis onset.

Materials and methods

Datasets

The data used in this study were obtained from a public domain database, which consisted of Intensive Care Unit (ICU) patient records in Beth Israel Deaconess Medical Center and Emory University Hospital, including a total of 40,336 patient records, collected over the past decade with approval from the appropriate Institutional Review Boards [28]. Each record consisted of a combination

of hourly vital sign summaries, laboratory values, and demographic variables. Specifically, the data contained 40 clinical variables: 8 vital sign variables, 26 laboratory variables, and 6 demographic variables. Tables 1 and 2 present these variables. We changed the definition of SepsisLabel to correspond to our experiment. A summary of vital signs and laboratory value data in the dataset is shown in Table 3. The missing rate of vital signs ranged from 9.88% (for the heart rate) to 66.16% (for temperature). Moreover, the missing rate of laboratory data ranged from 82.89% (for glucose) to 99.81% (for direct bilirubin). More details of the dataset are provided in a previous study [29].

Definition of sepsis onset time

We labeled patient data in accordance with the clinical criteria of Third International Consensus Definitions for Sepsis and Septic Shock. For each sepsis patient, we specified the following three time points to define the onset time t_{sepsis} of sepsis:

- $t_{suspicion}$: Clinical suspicion of infection identified as the earlier timestamp of intravenous (IV) antibiotics and blood cultures within a given time interval. If IV antibiotics were given first, then the cultures must have been obtained within 24 h. If cultures were obtained first, then IV antibiotics must have been ordered within 72 h. In either case, IV antibiotics must have been administered for at least 72 consecutive hours.

Table 2 Clinical time series data: laboratory values

	Measurement	Description
1	EtCO ₂	End tidal carbon dioxide (mm Hg)
2	BaseExcess	Excess bicarbonate (mmol/L)
3	HCO ₃	Bicarbonate (mmol/L)
4	FiO ₂	Fraction of inspired oxygen (%)
5	pH	pH
6	PaCO ₂	Partial pressure of carbon dioxide from arterial blood (mm Hg)
7	SaO ₂	Oxygen saturation from arterial blood (%)
8	AST	Aspartate transaminase (IU/L)
9	BUN	Blood urea nitrogen (mg/dL)
10	Alkalinephos	Alkaline phosphatase (IU/L)
11	Calcium	Calcium (mg/dL)
12	Chloride	Chloride (mmol/L)
13	Creatinine	Creatinine (mg/dL)
14	Bilirubin direct	Direct bilirubin (mg/dL)
15	Glucose	Serum glucose (mg/dL)
16	Lactate	Lactic acid (mg/dL)
17	Magnesium	Magnesium (mmol/dL)
18	Phosphate	Phosphate (mg/dL)
19	Potassium	Potassium (mmol/L)
20	Bilirubin total	Total bilirubin (mg/dL)
21	TroponinI	Troponin I (ng/mL)
22	Hct	Hematocrit (%)
23	Hgb	Hemoglobin (g/dL)
24	PTT	Partial thromboplastin time (s)
25	WBC	Leukocyte count (count/L)
26	Fibrinogen	Fibrinogen concentration (mg/dL)
27	Platelets	Platelet count (count/mL)

Table 1 Clinical time series data: vital signs (rows 1–7), demographics (rows 8–13), and outcome (row 14)

	Measurement	Description
1	HR	Heart rate (beats per minute)
2	O ₂ Sat	Pulse oximetry (%)
3	Temp	Temperature (°C)
4	SBP	Systolic BP (mm Hg)
5	MAP	Mean arterial pressure (mm Hg)
6	DBP	Diastolic BP (mm Hg)
7	Resp	Respiration rate (breaths per minute)
8	Age	Age (years)
9	Gender	Female (0) or male (1)
10	Unit1	Administrative identifier for ICU unit (MICU); false (0) or true (1)
11	Unit2	Administrative identifier for ICU unit (SICU); false (0) or true (1)
12	HospAdmTime	Time between hospital and ICU admission (hours since ICU admission)
13	ICULOS	ICU length of stay (hours since ICU admission)
14	SepsisLabel*	For septic patients, SepsisLabel is 1 if $t \geq t_{sepsis}$ and 0 if $t < t_{sepsis}$ For non-septic patients, SepsisLabel is 0

*We changed the definition of SepsisLabel to correspond to our experiment. The original definition is as follows: For septic patients, SepsisLabel is 1 if $t \geq t_{sepsis} - 6$ and 0 if $t < t_{sepsis} - 6$. For nonsepsis patients, SepsisLabel is 0

Table 3 Summary of vital signs and laboratory data in the datasets

Number of patients	40,336
Number of septic patients	2932
Sepsis prevalence	7.26%
Number of rows	1,424,147
Number of entries	10,486,913
Density of entries	19.9%

- t_{SOFA} : Occurrence of organ failure as identified by a 2-point increase in the Sequential Organ Failure Assessment (SOFA) score within a 24-h period.
- t_{sepsis} : Onset of sepsis identified as being earlier than $t_{\text{suspicion}}$ and t_{SOFA} as long as t_{SOFA} occurred no more than 24 h before or 12 h after $t_{\text{suspicion}}$

Data preprocessing

Missing values in the original data were intentionally preserved for conforming to the actual clinic situation. The missing values do become a challenge. Jang-Sikchoi's team used logistic regression as the algorithm, and last observation carried forward and K-nearest neighbors as imputation methods for sepsis screening [23]. Ujjwol' s team used gradient boosting tree as the algorithm and mean value as an imputation method for sepsis prediction [30]. To address this problem and to simulate the clinical situation, we imputed the missing values first by last observation carried forward. If the initial hourly data were unavailable for a variable in the patient record, the missing value was imputed with mean value calculated from the data for the variable in all 40,336 patient records. In addition to the original 40 variables, three new variables were created as follows: heart rate/systolic blood pressure, blood urea nitrogen/creatinine, and oxygen saturation from arterial blood/fraction of inspired oxygen.

Machine learning model

ANN was used as our machine learning model in this study. ANN pattern recognition was implemented using the MATLAB ANN toolbox, which was based on stochastic gradient descent (SGD). SGD is an iterative method for optimizing an objective function with suitable smoothness properties (e.g., differentiable or subdifferentiable). It can be regarded as a stochastic approximation of gradient descent optimization, because it replaces the actual gradient (calculated from the entire dataset) with an estimate thereof (calculated from a randomly selected subset of the data). Particularly, in high-dimensional optimization problems, this reduces the

computational burden, achieving faster iterations traded-off against a lower convergence rate.

We constructed a two-layer feed-forward network, with sigmoid hidden and softmax output neurons. The output layer was a softmax layer as an activation function outputting the probability of sepsis. The error function was evaluated based on cross-entropy and the percentage of misclassification errors.

The input of the classifier included all 43 variables: 8 vital sign variables, 26 laboratory variables, 6 demographic variables, and 3 created variables. To train our classifier, the number of hidden neurons was set as 100, 150 and 200 for one-hourly data; 300, 400 and 500 for three-hourly data; and 600, 700 and 800 for five-hourly data. The optimized number of hidden neurons was found by trial-and-error. Figure 1 provides the schematic of our ANN model.

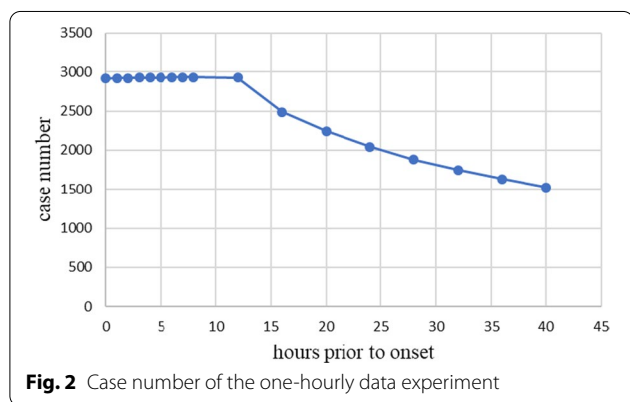
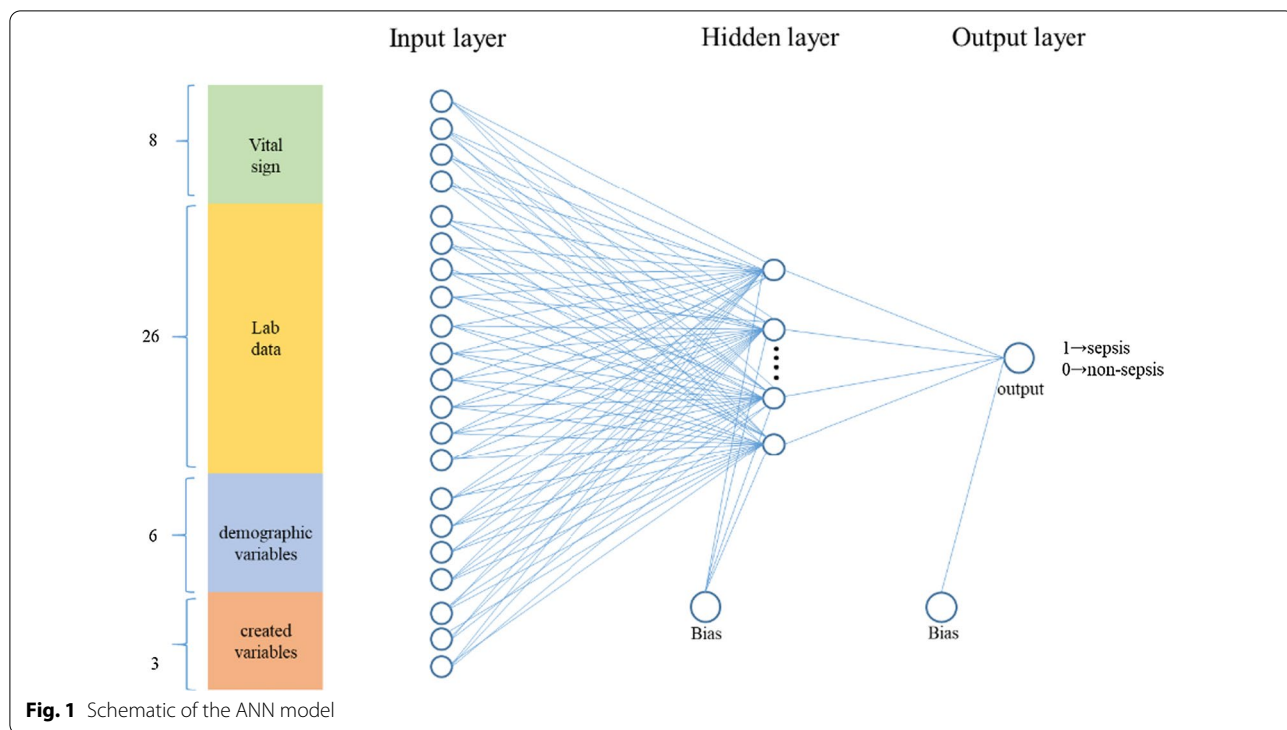
Experiment

Classifying sepsis and nonsepsis patients for predicting sepsis

In this study, 40,336 patients consisted of 2932 sepsis patients and 37,404 nonsepsis patients. When training artificial intelligence-based models with imbalanced data with significantly higher negative results than positive results, outcomes tend to be negative [31]. To address this problem, we adjusted the ratio of sepsis to nonsepsis patients to 1:1 by random matching. Next, to predict whether patients will develop sepsis, we extracted one-hourly data, three-hourly data, and five-hourly data prior to onset from all 2932 sepsis patients. How many hours prior to onset we would set depended on different experimental conditions. For example, in the one-hourly data experiment, the data of 1 h prior to sepsis onset were labeled as nonsepsis data initially in the dataset. We then defined the data to be sepsis data and used them to train our model to predict the status of sepsis 1 h in advance. We also randomly extracted one-hourly data, three-hourly data, and five-hourly data from 2932 nonsepsis patients, who were randomly matched to sepsis patients. These data were defined as nonsepsis data and used to predict the status of nonsepsis.

In the one-hourly data experiment, the number of hidden neurons was set as 200. We extracted one-hourly data over 0–40 h prior to onset separately from sepsis patient records. The details of case numbers are provided in Fig. 2. Each set of one-hourly data consisted of 43 variables. Thus, the number of inputs was 43.

In the three-hourly data experiment, the number of hidden neurons was set as 500. We extracted three-hourly data over 1–3 to 13–15 h prior to onset separately from sepsis patient records. The 1–3-h three-hourly sepsis data consisted of the data of 1 h,



2 h and 3 h prior to onset. Each set of three-hourly data consisted of 129 variables (3 × 43). Thus, the number of inputs was 129.

In the five-hourly data experiment, the number of hidden neurons was set as 800. We extracted five-hourly data over 1–5 to 16–20 h prior to onset separately from sepsis patient records. The 1–5-h five-hourly sepsis data consisted of the data of 1 h, 2 h, 3 h, 4 h and 5 h prior to onset. Each set of five-hourly data consisted of 215 variables (5 × 43). Thus, the number of inputs was 215.

Figure 3 shows the schematic of one-hourly, three-hourly, and five-hourly data experiments.

Using sepsis patient records only for predicting onset timing

We applied one-hourly data in this experiment, the number of hidden neurons was set as 200, and the number of inputs was 43. In sepsis patients, one-hourly data over 0–40 h prior to onset were extracted separately according to the experiment design; these data were termed as sepsis data. One-hourly data obtained prior to sepsis data were defined as nonsepsis data. We then adjusted the ratio of sepsis data to nonsepsis data to 1:1 by random matching. Figure 4 shows schematic of the experiment using only sepsis patient records.

Model validation and performance measurement

We divided the dataset into two groups: 85% of the data into a training group and 15% into a testing group, in order to build the 85% training and 15% testing cross-validation method. The training group was presented to the network during training, and the network was adjusted according to its error. Furthermore, 17.6% of the training group was used in algorithm of Levenberg–Marquardt to prevent over-fitting. The testing group had no effect on training and provided an independent measure of network performance after training. The training process ended when the gradient of performance was less than 10⁻⁶. Finally, we chose an adequate and well-trained model according to its training performance and testing performance.

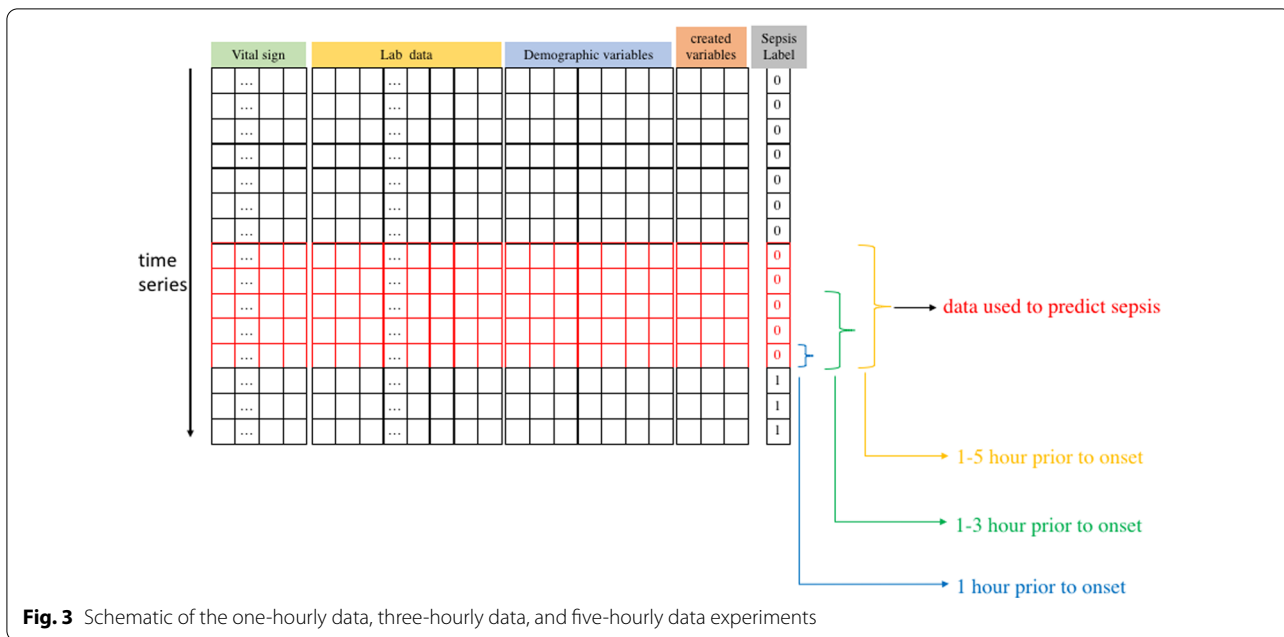


Fig. 3 Schematic of the one-hourly data, three-hourly data, and five-hourly data experiments

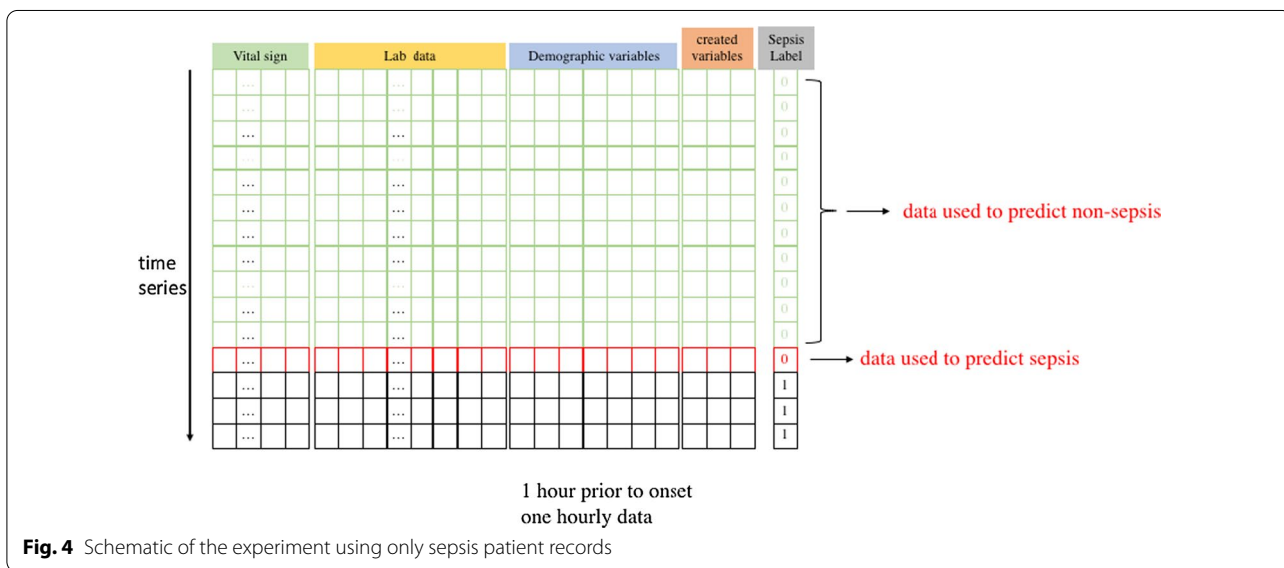


Fig. 4 Schematic of the experiment using only sepsis patient records

The model’s performance was determined using the area under the ROC curve (AUC) metric, sensitivity, and specificity. Sepsis and nonsepsis were set as positive and negative outcomes, respectively. We conducted all the experiments at a significance level of 95%.

Results

Classifying sepsis and nonsepsis patients for predicting sepsis

In each experiment, we trained 10 models with different random relative weights.

Performance in the one-hourly data experiment is shown in Table 4, Figs. 5 and 6. The AUC of the training group was the highest at 0 h prior to onset, which was used for identifying sepsis, and the mean AUC was 0.82. With an increase in the number of hours prior to onset, performance started to decline, reaching the lowest mean AUC 0.76 at 12 h prior to onset. No significant difference was found in the AUC performance of 0, 1 and 2 h prior to onset. Furthermore, the performance of more than 12 h prior to onset started to increase. A rebounding effect of the AUC performance was observed between 13

Table 4 Performance characteristics of the one-hourly data experiment

Hours prior to onset	AUC (95% CI)	Accuracy	Sensitivity	Specificity
0	0.821 (0.814–0.828)	0.758	0.836	0.678
6	0.779 (0.772–0.787)	0.731	0.793	0.669
12	0.759 (0.745–0.773)	0.703	0.772	0.633
24	0.791 (0.785–0.797)	0.738	0.817	0.657
36	0.807 (0.793–0.822)	0.742	0.793	0.690

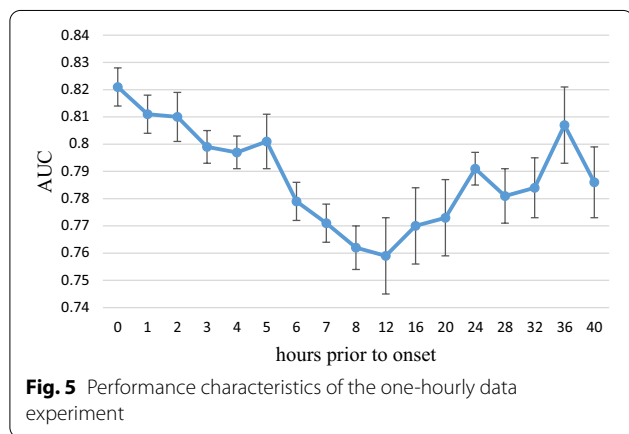
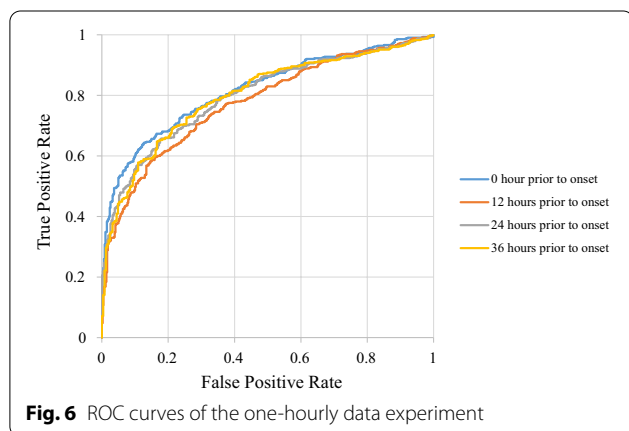
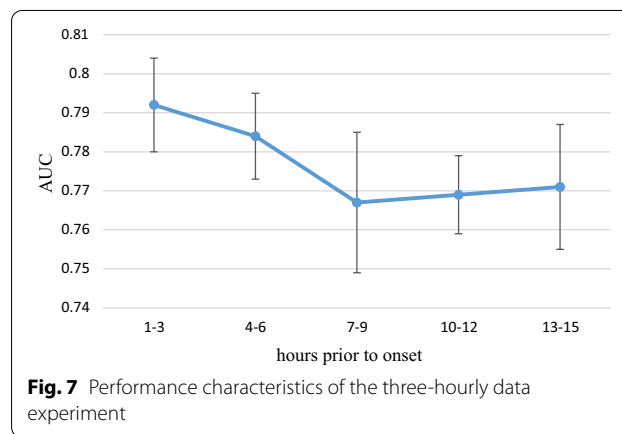


Table 5 Performance characteristics of the three-hourly data experiment

Hours prior to onset	AUC (95% CI)	Accuracy	Sensitivity	Specificity
1–3	0.792 (0.780–0.805)	0.730	0.796	0.665
4–6	0.784 (0.773–0.794)	0.721	0.777	0.665
7–9	0.767 (0.749–0.784)	0.703	0.754	0.654
10–12	0.769 (0.759–0.779)	0.711	0.762	0.660
13–15	0.771 (0.755–0.787)	0.716	0.775	0.642



and 40 h prior to onset. The mean AUC at 40 h prior to onset was 0.786. Thus, our ANN model has the moderate ability to predict whether patients will develop sepsis, even up to 40 h prior to sepsis onset. This ability can enable clinical health professionals to take appropriate measures beforehand to treat sepsis.

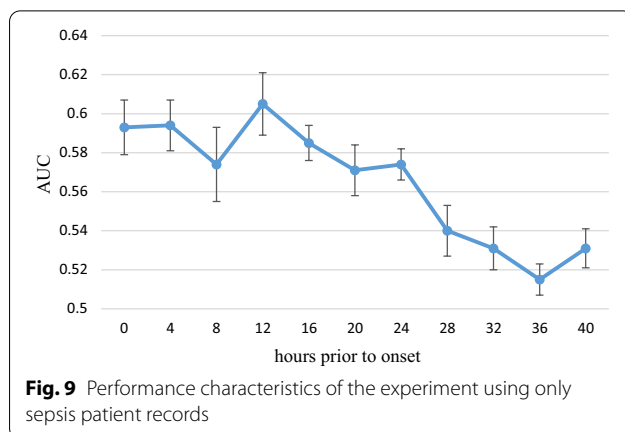
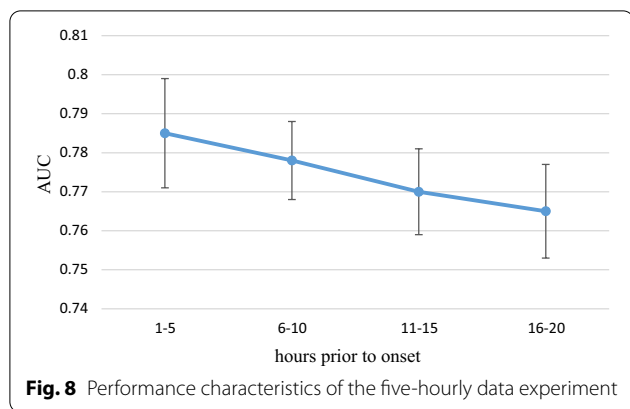
Performance in the three-hourly data experiment is shown in Table 5 and Fig. 7. The AUC of the training group was the highest at 1–3 h prior to onset, with a

Table 6 Performance characteristics of the five-hourly data experiment

Hours prior to onset	AUC (95% CI)	Accuracy	Sensitivity	Specificity
1–5	0.785 (0.771–0.800)	0.719	0.762	0.676
6–10	0.778 (0.768–0.788)	0.716	0.764	0.668
11–15	0.770 (0.759–0.781)	0.713	0.775	0.652
16–20	0.765 (0.753–0.778)	0.713	0.758	0.668

mean AUC of 0.792. Compared with the one-hourly data experiment, a higher performance was not found for the three-hourly data experiment. Although the AUC performance was the lowest (with mean AUC of 0.767) at 7–9 h prior to onset and a rebounding effect of the AUC performance was observed between 7–9 and 13–15 h prior to onset, no significant difference was found for the AUC performance of all three-hourly data experiments.

Performance in the five-hourly data experiment is presented in Table 6 and Fig. 8. The AUC of the training group was the highest at 1–5 h prior to onset, with a mean AUC of 0.785. Compared with the one-hourly data experiment and three-hourly data experiment, a



higher performance was not found for the five-hourly data experiment. The rebounding effect of AUC was not observed. With an increase in the number of hours prior to onset, performance kept declining, reaching the lowest mean AUC of 0.765 at 16–20 h prior to onset. No significant difference was observed in the AUC performance of all five-hourly data experiments.

Overall, the highest performance was found for one-hourly data experiment for identifying and predicting sepsis, demonstrating that our ANN model can perform adequately with limited hourly data provided.

Using sepsis patient records only for predicting onset timing

The performance of the experiment using only sepsis patient records is shown in Table 7 and Fig. 9. The mean AUC of the testing group ranged between 0.605 and 0.515. Compared with the experiment for classifying sepsis and nonsepsis patients, the performance of the experiment using only sepsis patient records was much lower, demonstrating that our ANN model is not suitable for precisely predicting the onset timing of sepsis or classifying the status of the same patient at different time point. Furthermore, no significant difference was found in the results at 0–24 h prior to onset, nor were any differences

Table 7 Performance characteristics of the experiment using only sepsis patient records

Hours prior to onset	AUC (95% CI)	Accuracy	Sensitivity	Specificity
0	0.593 (0.579–0.607)	0.572	0.572	0.548
4	0.594 (0.581–0.607)	0.573	0.590	0.556
8	0.574 (0.555–0.594)	0.563	0.587	0.539
12	0.605 (0.589–0.621)	0.586	0.611	0.561
24	0.574 (0.566–0.582)	0.548	0.557	0.540
36	0.515 (0.503–0.527)	0.523	0.517	0.529

found in the results at 28–40 h prior to onset. However, the results at 0–24 h prior to onset significantly outperformed the results at 28–40 h prior to onset, demonstrating that the data closer to sepsis onset had more predictive value.

Discussion

Classifying sepsis and nonsepsis patients for predicting sepsis

In clinical situations, it is unlikely that the complete data of every hour would be available. Missing and erroneous data are unavoidable under normal circumstances. Nonetheless, our ANN model showed a performance above 0.8 AUC in the presence of up to 80% missing and erroneous data, showing its ability of clinical application and proving that the model can predict whether patients will develop sepsis before sepsis onset and before significant changes in vital signs and laboratory data.

In our experiment, we could even predict whether patients would develop sepsis up to 40 h in advance prior to sepsis onset, with a performance of 0.786 AUC. Although this ANN model cannot precisely predict sepsis onset, it can identify patients who will develop sepsis 40 h in advance, which is valuable information for clinical and medical professionals. Therefore, they can provide adequate management and treatment 40 h in advance, including early source control, fluid therapy, vasoactive medications, and antibiotic administration [32]. According to some autopsy studies in adults, the most common error in the treatment of sepsis is the delay in diagnosing sepsis and infection treatment; this delay is avoidable if we are aware of the sepsis status of the patient in advance [33, 34]. The Surviving Sepsis Campaign (SCC), a joint collaboration between the European Society of Intensive Care Medicine, International Sepsis Forum, and the Society of Critical Care Medicine, has also emphasized the importance of early source control and antibiotic

administration [35]. Furthermore, SCC has shown that compliance with adequate early resuscitation and management bundle could significantly reduce sepsis mortality in hospitals [36]. Many studies have shown the benefits and advantages of early medical intervention for sepsis, with the early identification of sepsis. Therefore, our ANN model can be applied in clinical settings to provide sepsis onset prediction for clinical and medical professionals.

In the one-hourly data, three-hourly data, and five-hourly data experiments, we found adequate performance of the one-hourly data for classifying sepsis and nonsepsis patients in advance. The AUC performance of the one-hourly data experiment from 1 to 6 h prior to onset was between 0.797 and 0.811. The AUC performance of the three-hourly data experiment from 1 to 6 h prior to onset was between 0.784 and 0.792. The AUC performance of the five-hourly data experiment from 1 to 5 h prior to onset was 0.785. More hourly data as the input did not increase the performance of the model. Therefore, our ANN model only needs the initial one-hourly data, demonstrating that we can assess the sepsis risk of a patient with the initial vital signs and laboratory data.

Other studies have used machine learning models to classify sepsis patients and nonsepsis patients. Qingqing's team used gradient tree boosting as the algorithm and vital signs as the input. They observed a mean AUC of 0.90 at 0 h prior to onset at a 0% missing rate and a mean AUC of 0.75 at 0 h prior to onset at a 60% missing rate [21]. Christopher's team used gradient tree boosting as the algorithm and vital signs as the input. At a 0% missing rate, they observed a mean AUC of 0.88 at 0 h prior to onset, mean AUC of 0.84 at 24 h prior to onset, and mean AUC of 0.83 at 48 h prior to onset. However, the case number was only 375 and 147 at 24 and 48 h prior to onset, respectively [22]. Jang-Sikchoi's team used logistic regression as the algorithm and laboratory data as the input. They observed a mean AUC of 0.83 at 0 h prior to onset at a 7% missing rate [23]. Compared with the models in these studies, our ANN model provides more advantages in clinical situations, as our model was trained with data with an 80% missing rate and imputed under clinical situation.

Using sepsis patient records only for predicting onset timing

In this experiment, we used our ANN model to classify every hourly dataset of sepsis patients. We aimed to find out whether any significant differences exist in vital signs and laboratory data before sepsis onset, which we could use to precisely predict the timing of sepsis onset. However, favorable performance was not found in this

experiment, with the highest mean AUC of 0.6. Even though we tried to classify the hourly data when sepsis occurred, the mean AUC reached only 0.593, demonstrating that our model is not suitable for classifying every hourly dataset of sepsis patient. Therefore, our ANN model is not suitable for precisely predicting the timing of sepsis onset. An algorithm consisting of time series might be considered to build a model to predict the precise timing of sepsis onset.

Conclusions

In the experiment using sepsis patient and nonsepsis patient records, the mean AUC reached 0.821. Our ANN model has the moderate ability to predict whether patients will develop sepsis, even up to 40 h prior to sepsis onset under simulated clinical situation with real-world data. In addition, this might imply the presence of a significant difference between sepsis patients and nonsepsis patients, even at 40 h prior to sepsis onset. Nonetheless, in sepsis patients, regardless of how many hours prior to onset, a significant difference was not found in vital signs and laboratory data. This might have resulted in the poor performance of our ANN model.

The results showed the effectiveness of our ANN model for early classifying sepsis and nonsepsis patient. However, the predictive performance still needed to be improved. We hope to cope with this issue by optimizing the models, using novel imputation methods and pursuing new features closely related to sepsis such as monocyte distribution width [37]. In our ANN model, we have demonstrated that given one-hourly input data can identify and predict sepsis and the accuracy is comparable to given three-hourly and five-hourly input data, which need extra information from the patients, and the necessity of more hourly data as input will be further investigated in the future.

Limitation

With an increase in the number of hours prior to onset, the case number would decrease because some patient records would not include the hourly data that were long time before the onset of sepsis. It was unclear whether this would affect our results. The patient records including hourly data long time before the onset of sepsis may have more similar patterns, creating difficulty in evaluating the prediction performance of our ANN model.

To validate the mean and last observation carried forward method and perform the noise tolerance capability of our ANN model, an experiment using the dataset with no missing or erroneous values should be performed. However, the dataset consisting of laboratory data from blood test is difficult to have no missing value in clinical situation. Therefore, there is no hourly data consisting of

no missing values for all variables before imputation in our dataset, and we cannot perform this experiment.

In comparison to other machine learning models, the “black box” nature of ANNs acts as a barrier in providing biological interpretation of the model. We can hardly present the value that the variables provide, relation between variables and results, and the threshold of making a decision. Furthermore, ANN needs more data for training, for it consists of many hidden neurons, which means that more parameters are needed to figure out.

Authors' contributions

Y-YK and H-WC conceived of the presented idea and developed the methods. Y-YK carried out the experiment, built the models, wrote the manuscript and prepared all figures. S-TH provided the clinical insights. H-WC supervised the project. All authors discussed the results, contributed to the final manuscript and reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by Ministry of Science and Technology, Taiwan, R.O.C. under Grant Nos. MOST 109-2221-E-038-011 and MOST 110-2221-E-038-006.

Availability of data and materials

All patient records files are available from the PhysioNet Computing in Cardiology Challenge 2019 (<https://doi.org/10.13026/v64v-d857>).

Declarations

Ethics approval and consent to participate

The data used in this study were obtained from a public domain database, which consisted of Intensive Care Unit (ICU) patient records in Beth Israel Deaconess Medical Center and Emory University Hospital, including a total of 40,336 patient records, collected over the past decade with approval from Emory University Institutional Review Board (belongs to Emory University) approved protocol No. 33,069.

Consent for publication

This material has not been published in whole or in part elsewhere. All authors listed on the title page have read the manuscript, attest to the validity and legitimacy of the data and its interpretation, and agree to its submission to BMC Medical Informatics and Decision.

Competing interests

The Authors declare that there is no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Author details

¹School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan. ²Department of Emergency Medicine, Mackay Memorial Hospital, Taipei, Taiwan. ³Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan.

Received: 2 July 2021 Accepted: 12 October 2021

Published online: 22 October 2021

References

- Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA*. 2016;315(8):801–10. <https://doi.org/10.1001/jama.2016.0287>.
- Martin GS, Mannino DM, Eaton St, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. *N Engl J Med*. 2003;348:1546–54. <https://doi.org/10.1056/NEJMoa022139>.
- Kaukonen K-M, Bailey M, Suzuki S, Pilcher D, Bellomo R. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012. *JAMA*. 2014;311:1308–16. <https://doi.org/10.1001/jama.2014.2637>.
- Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit Care Med*. 2001;29:1303–10. <https://doi.org/10.1097/00003246-200107000-00002>.
- Angus DC, Kelley MA, Schmitz RJ, White A, Popovich, Jr J, for the Committee on Manpower for Pulmonary and Critical Care Societies (COMPACCS). Current and Projected Workforce Requirements for Care of the Critically Ill and Patients With Pulmonary Disease: Can We Meet the Requirements of an Aging Population? *JAMA*. 2000; 284(21):2762–2770. <https://doi.org/10.1001/jama.284.21.2762>.
- Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, Suppes R, Feinstein D, Zanotti S, Taiberg L, Gurka D, Kumar A, Cheang M. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*. 2006;34:1589–96. <https://doi.org/10.1097/01.CCM.0000217961.75225.E9>.
- Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet*. 2020;395(10219):200–11. [https://doi.org/10.1016/S0140-6736\(19\)32989-7](https://doi.org/10.1016/S0140-6736(19)32989-7).
- Berger T, Birnbaum A, Bijur P, Kuperman G, Gennis P. A computerized alert screening for severe sepsis in emergency department patients increases lactate testing but does not improve inpatient mortality. *Appl Clin Inform*. 2010;1:394–407. <https://doi.org/10.4338/ACI-2010-09-RA-0054>.
- Hooper MH, Weavind L, Wheeler AP, Martin JB, Gowda SS, Semler MW, Hayes RM, Albert DW, Deane NB, Nian H, Mathe JL, Nadas A, Sztipanovits J, Miller A, Bernard GR, Rice TW. Randomized trial of automated, electronic monitoring to facilitate early detection of sepsis in the intensive care unit. *Crit Care Med*. 2012;40:2096–101. <https://doi.org/10.1097/CCM.0b013e318250a887>.
- Semler MW, Weavind L, Hooper MH, Rice TW, Gowda SS, Nadas A, Song Y, Martin JB, Bernard GR, Wheeler AP. An electronic tool for the evaluation and treatment of sepsis in the ICU: a randomized controlled trial. *Crit Care Med*. 2015;43:1595–602. <https://doi.org/10.1097/CCM.000000000000101020>.
- Seetharaman S, Wilson C, Landrum M, Qasba S, Katz M, Ladikos N, Harris JE, Galitsatos P, Yousem DM, Knight AM, Pearse DB, Blanding R, Bennett R, Galai N, Perl TM, Sood G. Does use of electronic alerts for systemic inflammatory response syndrome (SIRS) to identify patients with sepsis improve mortality? *Am J Med*. 2019;132:862–8. <https://doi.org/10.1016/j.amjmed.2019.01.032>.
- Chakraborty RK, Burns B. Systemic inflammatory response syndrome. [Updated 2020 Apr 28]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK547669/>.
- Simeone O. A very brief introduction to machine learning with applications to communication systems. *IEEE*. 2018;4:648–64. <https://doi.org/10.1109/TCCN.2018.2881442>.
- Reismann J, Romualdi A, Kiss N, Minderjahn MI, Kallarackal J, Schad M, Reismann M. Diagnosis and classification of pediatric acute appendicitis by artificial intelligence methods: an investigator-independent approach. *PLOS ONE*. 2019;14: e0222030. <https://doi.org/10.1371/journal.pone.0222030>.
- Kavakiotisab I, Tsavec O, Salifoglouc A, Maglaverasbd N, Vlahavasa I, Chouvardabd I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. 2017;15:104–16. <https://doi.org/10.1016/j.csbj.2016.12.005>.
- Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2015;13:8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>.
- Mohammed A, Van Wyk F, Chinthala LK, Khojandi A, Davis RL, Cooper-smith CM, Kamaleswaran R. Temporal differential expression of physi-omarkers predicts sepsis in critically ill adults. *Shock*. 2021;56(1):58–64. <https://doi.org/10.1097/SHK.0000000000001670>.

18. Futoma J, Simons M, Doshi-Velez F, Kamaleswaran R. Generalization in clinical prediction models: the blessing and curse of measurement indicator variables. *Crit Care Explor.* 2021;3(7): e0453. <https://doi.org/10.1097/CCE.0000000000000453>.
19. Wong AI, Cheung PC, Kamaleswaran R, Martin GS, Holder AL. Machine learning methods to predict acute respiratory failure and acute respiratory distress syndrome. *Front Big Data.* 2020;3: 579774. <https://doi.org/10.3389/fdata.2020.579774>.
20. Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med.* 2020;46(3):383–400. <https://doi.org/10.1007/s00134-019-05872-y>.
21. Mao Q, Jay M, Hoffman JL, Calvert J, Barton C, Shimabukuro D, Shieh L, Chettipally U, Fletcher G, Kerem Y, Zhou Y, Das R. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department. *General Ward and ICU BMJ Open.* 2018;26:e017833. <https://doi.org/10.1136/bmjopen-2017-017833>.
22. Barton C, Chettipally U, Zhou Y, Jiang Z, Lynn-Palevsky A, Le S, Calvert J, Das R. Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. *Comput Biol Med.* 2019;109:79–84. <https://doi.org/10.1016/j.combiomed.2019.04.027>.
23. Choi JS, Trinh TX, Ha J, et al. Implementation of complementary model using optimal combination of hematological parameters for sepsis screening in patients with fever. *Sci Rep.* 2020;10(1):273. <https://doi.org/10.1038/s41598-019-57107-1>.
24. Scarborough D, Somers M. Neural networks in organizational research: applying pattern recognition to the analysis or organizational behavior. *Am Psychol Assoc.* 2006. <https://doi.org/10.1037/11465-000>.
25. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: a survey. *Heliyon.* 2018;4(11): e00938. <https://doi.org/10.1016/j.heliyon.2018.e00938>.
26. Lee KH, Dong JJ, Jeong SJ, Chae M-H, Lee BS, Kim HJ, Ko SH, Song YG. Early detection of bacteraemia using ten clinical variables with an artificial neural network approach. *J Clin Med.* 2019;8:1592. <https://doi.org/10.3390/jcm8101592>.
27. Helguera-Repetto AC, Soto-Ramírez MD, Villavicencio-Carrisoza O, et al. Neonatal sepsis diagnosis decision-making based on artificial neural networks. *Front Pediatr.* 2020;8:525. <https://doi.org/10.3389/fped.2020.00525>.
28. M. A. Reyna et al. Early prediction of sepsis from clinical data: the PhysioNet/computing in cardiology challenge 2019. In: 2019 Computing in cardiology. Singapore; 2019, p. 1–4. <https://doi.org/10.23919/CinC49843.2019.9005736>.
29. Reyna MA, Josef CS, Jeter R, Shashikumar SP, Westover MB, Nemati S, Clifford GD, Sharma A. Early prediction of sepsis from clinical data: the PhysioNet/computing in cardiology challenge. *Crit Care Med.* 2019;48:210–7.
30. Shrestha U, Alsadoon A, Prasad PWC, et al. Supervised machine learning for early predicting the sepsis patient: modified mean imputation and modified chi-square feature selection. *Multimed Tools Appl.* 2021;80:20477–500. <https://doi.org/10.1007/s11042-021-10725-2>.
31. He H, Garcia EA. Learning from imbalanced data. *IEEE.* 2009;21:1263–84. <https://doi.org/10.1109/TKDE.2008.239>.
32. Rhodes A, Evans LE, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med.* 2017;43:304–77. <https://doi.org/10.1007/s00134-017-4683-6>.
33. Gurnani PK, Patel GP, Crank CW, et al. Impact of the implementation of a sepsis protocol for the management of fluid-refractory septic shock: a single-center, before-and-after study. *Clin Ther.* 2010;32(7):1285–93. <https://doi.org/10.1016/j.clinthera.2010.07.003>.
34. Rehmani RS, Memon JI, Al-Gammal A. Implementing a collaborative sepsis protocol on the time to antibiotics in an emergency department of a Saudi hospital: quasi randomized study. *Crit Care Res Pract.* 2014;2014: 410430. <https://doi.org/10.1155/2014/410430>.
35. Marshall JC, Dellinger RP, Levy M. The surviving sepsis campaign: a history and a perspective. *Surg Infect (Larchmt).* 2010;11(3):275–81. <https://doi.org/10.1089/sur.2010.024>.
36. Levy MM, Dellinger RP, Townsend SR, et al. The surviving sepsis campaign: results of an international guideline-based performance improvement program targeting severe sepsis. *Intensive Care Med.* 2010;36(2):222–31. <https://doi.org/10.1007/s00134-009-1738-3>.
37. Crouser ED, Parrillo JE, Seymour CW, et al. Monocyte distribution width: a novel indicator of sepsis-2 and sepsis-3 in high-risk emergency department patients. *Crit Care Med.* 2019;47(8):1018–25. <https://doi.org/10.1097/CCM.0000000000003799>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

