A Minimal Set of Physiomarkers in High Frequency Real-Time Physiological Data Streams Predict Adult Sepsis Onset Earlier

Article *in* International Journal of Medical Informatics · December 2018

DOI: 10.1016/j.ijmedinf.2018.12.002

6 authors, including:

Franco van Wyk
University of Tennessee
7 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Anahita Khojandi
University of Tennessee
24 PUBLICATIONS **20** CITATIONS
SEE PROFILE

Akram Mohammed
University of Nebraska at Lincoln
47 PUBLICATIONS **44** CITATIONS
SEE PROFILE

Rishikesan Kamaleswaran
The University of Tennessee Health Science Center
30 PUBLICATIONS **59** CITATIONS
SEE PROFILE

International Journal of Medical Informatics 122 (2019) 55–62

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

A minimal set of physiomarkers in continuous high frequency data streams predict adult sepsis onset earlier

Franco van Wyk[a], Anahita Khojandi[a], Akram Mohammed[b], Edmon Begoli[a,c], Robert L. Davis[b], Rishikesan Kamaleswaran[b,*]

[a] University of Tennessee, Knoxville, TN, USA
[b] Center for Biomedical Informatics, Department of Pediatrics, University of Tennessee Health, USA Science Center, Memphis, TN, USA
[c] Oak Ridge National Laboratory, Knoxville, TN, USA

ARTICLE INFO

Keywords:
Sepsis
Physiological data
Artificial intelligence
Predictive model
Critical care

ABSTRACT

Purpose: Sepsis is a life-threatening condition with high mortality rates and expensive treatment costs. To improve short- and long-term outcomes, it is critical to detect at-risk sepsis patients at an early stage.
Methods: A data-set consisting of high-frequency physiological data from 1161 critically ill patients was analyzed. 377 patients had developed sepsis, and had data at least 3 h prior to the onset of sepsis. A random forest classifier was trained to discriminate between sepsis and non-sepsis patients in real-time using a total of 132 features extracted from a moving time-window. The model was trained on 80% of the patients and was tested on the remaining 20% of the patients, for two observational periods of lengths 3 and 6 h prior to onset.
Results: The model that used continuous physiological data alone resulted in sensitivity and F1 score of up to 80% and 67% one hour before sepsis onset. On average, these models were able to predict sepsis 294.19 ± 6.50 min (5 h) before the onset.
Conclusions: The use of machine learning algorithms on continuous streams of physiological data can allow for early identification of at-risk patients in real-time with high accuracy.

1. Introduction

In the critical care environment, the availability of vast volumes of data present a unique opportunity to generate novel insights for better care [1,2]. The analysis of substantial volumes of data are more tractable with the use of new sophisticated and efficient machine learning methods and strategies [1,3–5]. The management of sepsis can benefit from the use of such tools, specifically to identify at-risk patients earlier. Sepsis is a deadly life-threatening condition that arises from a significantly dysregulated response to infection, resulting in acute single or multi-organ failure and death [6,7]. If recognition is delayed, sepsis can rapidly progress to multiple organ dysfunction (MOD), resulting in high mortality rates [6,8]; an increase of approximately 8% in mortality rate is observed for each hour of delayed diagnosis of sepsis [9]. Predictive analytics applied to routinely collected continuous data, such as physiological data, can reduce recognition gaps while allowing for targeted and early goal-directed therapy, while improving situational awareness in critical care.

Machine learning techniques have been extensively used in medical decision making and treatment planning. For instance, these algorithms have been used to predict at-risk patients or patient outcomes, and to reduce alarm fatigue [10–12]. Similarly, machine learning algorithms have been successfully implemented in various medical image analyses to assist diagnosis and therapy in neurology, cardiology, and the detection of various cancers [13–20]. While, to date, machine learning algorithms have shown promise in detecting and predicting sepsis [21,22], much of the recent work has been centered around static and often manually entered electronic health record (EHR) data [23]. Recent work has shown that 'physiomarkers,' such as reduced heart rate variability, may precede the onset of sepsis [24–26], enabling a window of early recognition and treatment. In this paper, we utilize machine learning to predict the onset of sepsis in patients who are admitted to the intensive care unit (ICU), using continuous minute-by-minute data captured at the bedside.

1.1. Contributions

This paper introduces a novel method for applying a machine learning pipeline to high-frequency data streams in the area of sepsis prediction. Therefore, we make the following key contributions in the

* Corresponding author at: 50N. Dunlap Street, 491R, Memphis, TN, 38103, USA.
E-mail address: rkamales@uthsc.edu (R. Kamaleswaran).

https://doi.org/10.1016/j.ijmedinf.2018.12.002
Received 21 August 2018; Received in revised form 23 October 2018; Accepted 10 December 2018

area of precision medicine as applied to critical care medicine:

1 A predictive sepsis model built using a minimal set of six continuous, routinely collected bedside physiological data streams
2 An analysis pipeline tailored for 'online' implementation
3 Identification of salient physiomarkers that predict the onset of sepsis in critically ill adults using an integrated machine learning approach

2. Materials and methods

2.1. Data characterization

Continuous minute-by-minute physiological data was captured using a proprietary Cerner CareAware iBus® platform [27] at the Methodist LeBonheur Healthcare (MLH) System in Memphis TN between February and December 2017. The data was collected across Intensive Care Units (ICUs) in four adult hospitals within the MLH system. We captured heart rate (HR), diastolic blood pressure (DBP), systolic blood pressure (SBP) (via cuff if arterial not available), mean arterial pressure (MAP), temperature, respiratory rate (RR) and peripheral oxygen saturation (SpO_2) based on the availability of the data. In addition, the white blood cell (WBC) count was collected from patients' electronic health records. We defined the onset of sepsis by following the sepsis-2 definition: patients were defined as having sepsis when they met Systemic Inflammatory Response Syndrome (SIRS) criteria in the presence of the suspicion of infection, indicated by the presence of a blood culture and the administration of antibiotics during the encounter, along with relevant International Classification of Diseases, Tenth Revision (ICD10) [28] codes appearing in the primary and additional discharge columns (see Supplemental Tables 1 and 2) upon discharge. (While there is growing interest in applying the new sepsis-3 definition, and incorporating the sequential organ failure assessment score (SOFA), MLH [29] had yet to endorse the new definition, thus we aligned our definition to the sepsis-2 workflow.) For the purpose of this study, we excluded patients without at least 8 h of continuous data, or who had ICD 10 codes in their discharge summary that indicated cardiovascular disease, including but not restricted to congestive heart failure, arrhythmias, or myocardial infarction (see consort diagram in the Fig. 1). Those patients were excluded to make the study subjects more homogenous as many of those patients had been receiving medication for modulation of the heart rate, potentially impacting one of the primary physiological inputs used in the models. Separate models can be developed for those patients to more closely capture and learn from their particular heart rate patterns presented to predict sepsis. Institutional Research Board (IRB) approval was received for this retrospective observational study.

2.2. Features extracted from moving time windows

Fig. 2 presents an overview of our approach. Physiological data including HR, DBP, SBP, temperature, RR, and SpO_2 are continuously used to make predictions about sepsis onset. Specifically, a fixed-width moving time-window of data is used as an input to the model, where the output is the probability of developing sepsis. Each minute, new observations in the form of physiological data are made on the patient at the bedside. Hence, the time window moves forward to include the data corresponding to these new observations, while the earliest data are removed from the time-window to maintain its width. Once these input data are updated, a new prediction about the likelihood of the patient developing sepsis is made. Note that intermittent laboratory data such as WBC count may be used in such a model, where the last observation is carried forward.

We allow the time-window to span one hour. Each minute, physiological data from the moving time-window were used as a model input. These inputs included 48 parameters (basic statistics and signal information) describing the entire one-hour time interval in addition to 42 parameters (basic statistics) from the last 30 and 15 min of the data stream. The temporal segments were identified using a preliminary exploration of various durations to achieve optimal performance. The processed data were combined, making up a total of $(48 + (42 \times 2) =)$ 132 features, which served as the features of each one-hour time-window in the model that relies on continuous physiological data alone. In addition, we build a similar model, where we also include WBC count as a feature in the model. WBC is frequently used to diagnose (bacterial) sepsis; however, it is often time-delayed and subject to an influence from a wide array of clinical factors [30]. Therefore, in this paper we assess the predictive role of WBC when it is used in conjunction with continuous physiological data streams to examine its (added) value. In the results section, we comment on the comparison in performance of these two models, respectively referred to as (*i*) physiological-only data and (*ii*) physiological data with WBC. We analyzed data from one-hour

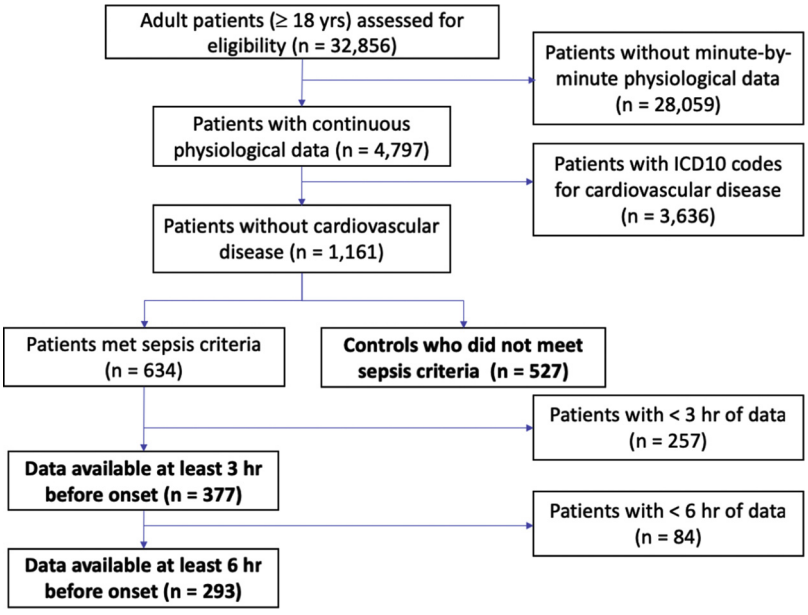

Fig. 1. Consort diagram of the patient population.

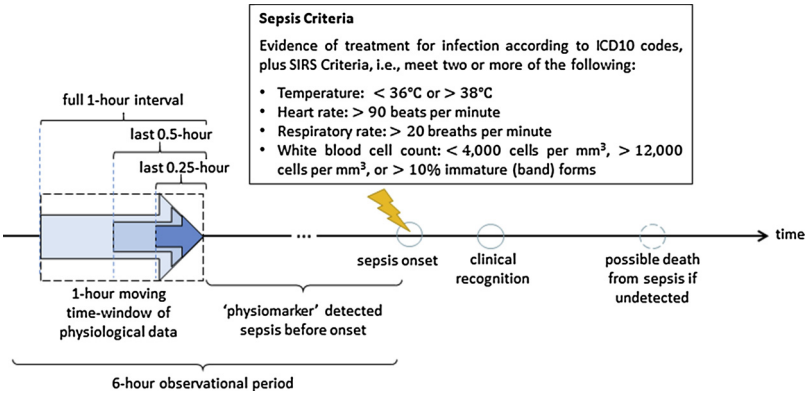

Fig. 2. An overview of the approach (with a 6-hour observational period).

moving time-windows for an overall observational period covering three and six hours and developed two machine learning models. As seen in Fig. 2, the models would start producing results, after the minimum data requirement of one hour is met. Longer observational periods were considered but were not pursued because of poor data availability.

2.3. Model development

The random forest (RF) algorithm is a supervised classification algorithm that relies on the aggregate classification or the 'majority vote' of a series of decision trees, each built using random subsets of the dataset [31]. Our prior experience using RF have generated state-of-the-art results for 'one-time' predictions using high-frequency datasets [21], which we aim to extend to 'online' predictions in this work. In general, RF algorithms outperform other tree-based classification algorithms such as decision tree learning and tree bagging [32,33]. Random forest algorithms are generally robust against overfitting and are not overly sensitive to noisy data or a high ratio of parameters to observations [34]. In addition, RF algorithms can be used to identify and rank the most important features contributing to classification performance. Lastly, compared to machine learning algorithms such as neural networks or other deep learning techniques, RF algorithms are known for their fast training and computation time for large datasets [35,36], making them appropriate in our application for conducting various experiments to determine best window sizes.

In this study, 80% of the data were used for training the models, with the goal of achieving the lowest out-of-bag (OOB) error [31]. The two best resulting models for 3- and 6-hour observational periods were tested on the remaining 20% of the data to report their performance. In addition, the number of trees in the RF models, chosen based on the OOB, was set to 700 to assure that the classification error converges as a function of the number of trees for all models. Based on extensive experiments, data from the last 75 min in each observational period were used to train the RF algorithms. In this study, we predict sepsis to be present when the model-estimated probability of sepsis goes above 0.5. While our prior results have suggested the RF algorithm to be superior to other classifiers for our specific data, in this paper we present a comparative analysis of other methods, including Support Vector Machines (SVM), Logistic Regression (LR), Multi-Layer Perceptron (MLP) and Recurrent Neural Networks (RNN).

2.4. Performance metrics

To evaluate the model performance, we used sensitivity, specificity, positive predictive value (PPV), accuracy, F1 score, and area under the receiver operating characteristic curve (AUC) of the developed RF algorithms when applied to the test sets. Sensitivity measures the proportion of correctly predicted sepsis patients from the sepsis group. Similarly, specificity measures the proportion of correctly predicted non-sepsis patients among the non-sepsis group. Positive predictive value measures the proportion of sepsis patients among those predicted as positives. Accuracy gives the overall proportion of correct predictions. F1 score, which is another measure of a test accuracy, is the harmonic mean of sensitivity and PPV. Lastly, AUC is a measure of how well the algorithm can distinguish between the two diagnostic groups, i.e., sepsis and non-sepsis groups.

2.5. Implementation

The implementation of our methods was performed on a standalone server housed at the University of Tennessee Health Science Center Information Technology Services facility. In order to ensure generalizability of our model we developed our approach using open-source technology. Due to the high-frequency nature of our data, we used the Apache Spark [37] platform to preprocess and standardize our dataset. We used an existing Python Scikit-learn machine learning library [38,39] for conducting statistical analyses, feature selection and developing the machine learning classifiers. MLP and RNN classifiers were developed using Keras and Tensorflow [40,41]. We used these libraries due to their ability to enable an enhanced degree of hyper-parameterization at a high level within the open-source suite of platforms. All models were saved and versioned for reproducibility.

3. Results

Table 1 illustrates several descriptive statistics for the sepsis and non-sepsis patient subgroups for the 3-hour observational period, all of whom were aged between 18–99 years of age. A *t*-test between the sepsis and non-sepsis subgroups indicate statistically significant differences across all parameters based on p-values.

Balanced training and test sets, with an equal number of sepsis and

Table 1
Descriptive statistics for the sepsis and non-sepsis patient subgroups for the 3-hour observational period. The mean and 95% confidence intervals are presented in the table.

Parameter	Sepsis (n = 377)	Non-sepsis (n = 377)
HR (bpm)	84.86 ± 0.12	79.25 ± 0.11
RR (breaths per minute)	17.94 ± 0.03	17.56 ± 0.02
SBP (mmHg)	128.22 ± 0.17	135.79 ± 0.18
DBP (mmHg)	70.55 ± 0.11	74.19 ± 0.13
Temperature (°C)	36.87 ± 0.01	36.80 ± 0.01
Peripheral oxygen saturation (SpO_2)	97.17 ± 0.02	96.66 ± 0.04
WBC (10^9/L)	11.21 ± 0.04	9.04 ± 0.06
LOS (days)	11.72 ± 0.89	6.64 ± 0.57

HR: heart rate; RR: respiratory rate; SBP: systolic blood pressure; DBP: diastolic blood pressure; WBC: white blood cell; LOS: length of stay.

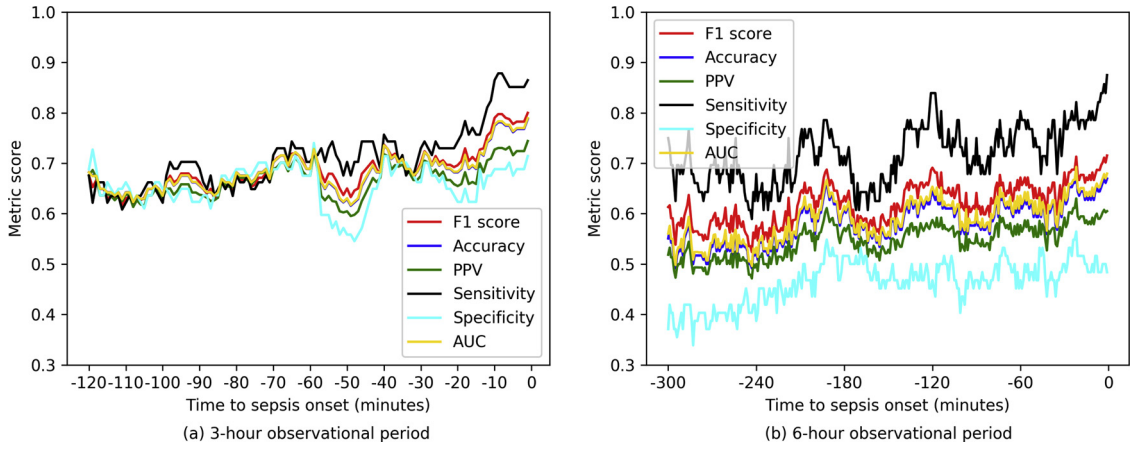

Fig. 3. F1 score, accuracy, PPV, sensitivity, specificity, and AUC of the trained RF classifiers in discriminating sepsis and non-sepsis patients when used in an online fashion on separate test sets for 3- and 6-hour observational periods.

non-sepsis patients, were generated to avoid favoring the more represented observations in the dataset. As discussed in Section 2.2, observational periods of lengths three and six hours were used. For instance, for the sepsis case group, we selected a 3-hour observational period prior to sepsis onset (as defined by SIRS criteria), and in the control group, which consists of non-sepsis patients, we used a given 3-hour observational period. Hence, at the end of the observational period, half of the patients developed sepsis. A total of 377 and 293 sepsis patients were included for the 3- and 6-hour observational periods, respectively (see Fig. 1).

Fig. 3 presents the F1 score, accuracy, PPV, sensitivity, specificity, and AUC of the RF algorithms for the physiological data-only model when applied to the test set in a real-time fashion for 3- and 6-hour observational periods. These metrics are calculated every minute during the observational period, beginning after the initial one-hour buffer for feature extraction, hence providing an aggregated snap-shot view of the performance of the RF algorithm as a function of time to sepsis onset. Within both observational window scenarios, predictions are generated from 1-hour feature extraction periods that are incrementally shifted in time, therefore, for instance, as illustrated in Fig. 3(a), the first prediction is available at the -120th minute for the 3-hour observational window. In Fig. 3(a), F1 score, accuracy, sensitivity, and AUC generally increase closer to the onset of sepsis while PPV and specificity stay relatively constant, ranging between 55%–74%. For instance, as seen in Fig. 3(a), the algorithm developed using the 3-hour observational period can discriminate patients with F1 score and sensitivity of 69% and 70% an hour before the onset of sepsis. These increase to 70% and 73%, respectively, half an hour before the onset. Similarly, in Fig. 3(b), for the 6-hour observational period, F1 score, accuracy, sensitivity, and PPV increase 1–1.5 hours before sepsis onset while specificity slightly increases throughout the observational period. The algorithm can discriminate patients with F1 score and sensitivity of 67% and 80% an hour before sepsis onset, with F1 score slightly increasing to 68% and sensitivity staying constant at 80%, half an hour before the onset.

The results obtained from the physiological data-only model, presented in Fig. 3, are similar in structure to those obtained from a model that relies on physiological data with WBC. The algorithm developed using the 3-hour and 6-hour observational period can discriminate patients with F1 scores of 71% and 69%, and sensitivity of 68% and 80%, respectively, an hour before the onset of sepsis. The increase in F1 scores, compared to the physiological data-only model with the corresponding F1 scores of 69% and 67%, is mainly due to the increase in specificity. For instance, for the 3-hour observational period, the specificity, which was up to 68% an hour before the onset of sepsis in the physiological data-only models, increases to up to 78% when also accounting for WBC. In addition to the RF model, we performed a comparative analysis using the same training and testing patient information to ensure a consistent comparison. Specifically, we use the same features as the RF model for the SVM, LR, and MLP models. For the RNN model, often used in time-series analysis, raw physiological data is fed into the network where automatic feature extraction is performed by the RNN. Supplemental Table 3 illustrates the performance comparison of the various classification algorithms for the 6 h observation period using physiological data. The results illustrate that RF outperforms the other algorithms in terms of sensitivity and F1 score at an hour and half an hour before sepsis onset, respectively. Fig. 4 presents the probability of developing sepsis for two example sepsis patients as estimated by the physiological data-only model in an online fashion during the 3-hour and 6-hour observational periods. As seen in Fig. 4(a), the probability of developing sepsis is consistently low and then increases approximately an hour and half before the sepsis onset. As seen in Fig. 4(a), for this example patient the algorithm would predict sepsis 93 min before onset (using the 0.5 cutoff), allowing for directing the attention of the clinical staff to the patient to possibly prevent sepsis development and improve patient outcomes. Similarly, as seen in Fig. 4(b), for this example patient the algorithm would predict sepsis 4 h and 50 min before onset. In general, when the physiological data-only model is able to predict sepsis, it does so on average (with a 95% confidence interval) 109.78 ± 5.27 min and

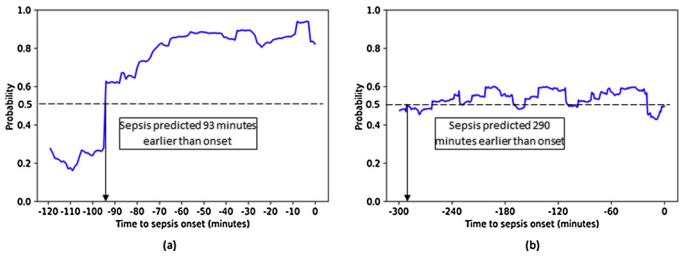

Fig. 4. The model-estimated probability of developing sepsis for two given sepsis patients over the 3-hour and 6-hour observational periods leading to sepsis onset.

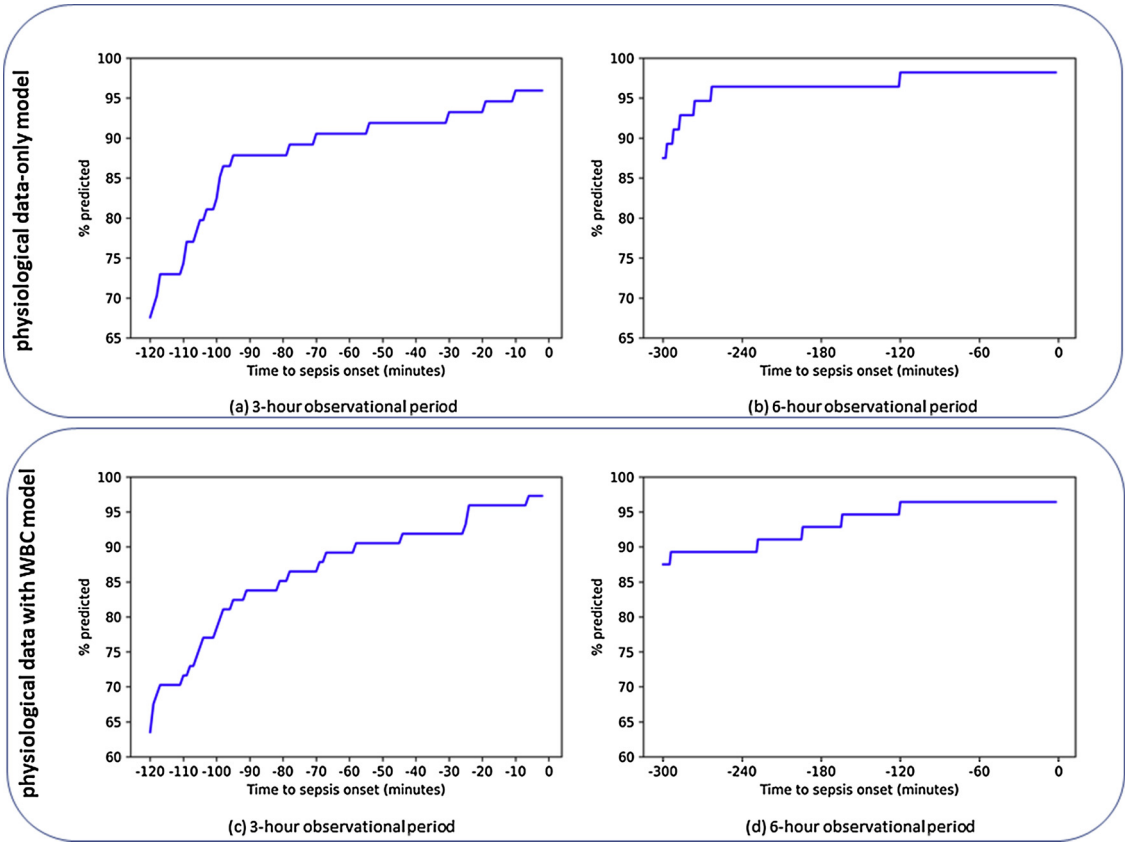

Fig. 5. Percentage of sepsis patients correctly predicted during the 3-hour and 6-hour observational periods, for the physiological data-only and physiological data with WBC models.

294.19 ± 6.50 min before the onset for the 3- and 6-hour observational periods, respectively. Compare these numbers with 106.28 ± 6.14 min and 289.75 ± 8.93 min, respectively, for the physiological data with WBC model.

Fig. 5 presents the percentage of sepsis patients correctly predicted during the 3-hour and 6-hour observational periods using the corresponding models. As seen in the figure, consistent with model performances with respect to the average lead times in predicting sepsis, in general the physiological data-only model predicts sepsis before the physiological data with WBC models. For instance, in Fig. 5(a), the physiological data-only model correctly predicts 88% of sepsis patients 90 min before onset for the 3-hour observational period. Compare this number with 84% for the physiological data with WBC model 90 min before onset for the 3-hour observational period.

Fig. 6 presents the aggregated mean and 95% confidence interval of the model-estimated probabilities for the test sets in a real-time fashion during the 3-hour observational period, stratified across the sepsis and non-sepsis subgroups. The top and bottom panels show the results for the physiological data-only and physiological data with WBC, models, respectively. As seen in Fig. 6(a), the probability of being classified as an at-risk sepsis patient increases for the sepsis subgroup during the observational period as the time of sepsis onset approaches. In contrast, as illustrated in Fig. 6(b), the mean false positive rate stays relatively constant between 0.36 and 0.43 during the entire observational period. In addition, the confidence interval width decreases for the sepsis subgroup as the time to sepsis onset approaches and physiomarkers become more apparent, illustrating more confidence in the model estimated probabilities. In contrast, the confidence interval width for the non-sepsis subgroup slightly increases as the time to sepsis onset approaches. These results are consistent with the ones presented in Fig. 6(c)–(d). However, as seen across the top and bottom panels, the addition of WBC as an input variable overall increases the model-estimated probabilities for the sepsis-positive class and decreases them for the sepsis-negative class. Hence, the addition of WBC overall results in more confident discrimination of patients at an aggregate level and a lower rate of false positives.

4. Discussion and conclusion

Continuous monitoring applied to the development of sepsis in hospitalized patients can reduce the gap between underlying pathological onset and clinical recognition. In this study, we follow the sepsis-2 definition, a widely used indicator of sepsis, to automatically identify the time of onset of sepsis. We further demonstrate that physiomarkers exist prior to the onset of sepsis, independent of WBC values. Furthermore, a unique advantage of our modelling approach is the minimal set of required variables, which only include six routinely collected physiological data (heart rate, respiratory rate, SBP, DBP, SpO_2, and temperature), in comparison to existing sepsis prediction models. The use of a minimal set of variables in our model ultimately makes it easier to implement at the bedside and without reliance on lagged data such as clinical observations or laboratory results.

In this paper, we present a method for online analysis of physiological data captured at the bedside to generate predictions about sepsis in all critically ill patients. Our results indicate that the continuous physiological data-only model is able to discriminate the case and control group with an average F1 score and sensitivity of up to 67% and 80% an hour before sepsis onset, and 68% and 80% half an hour before sepsis onset. The performance of the models developed for 3- and 6-hour observational periods are slightly different, mainly because of the difference in the number of patients included (see Fig. 1). On average, these models were able to predict sepsis 294.19 ± 6.50 min (5 h) before the onset. Interestingly, including WBC as a variable did not improve the model sensitivity, nor did it help with earlier prediction of

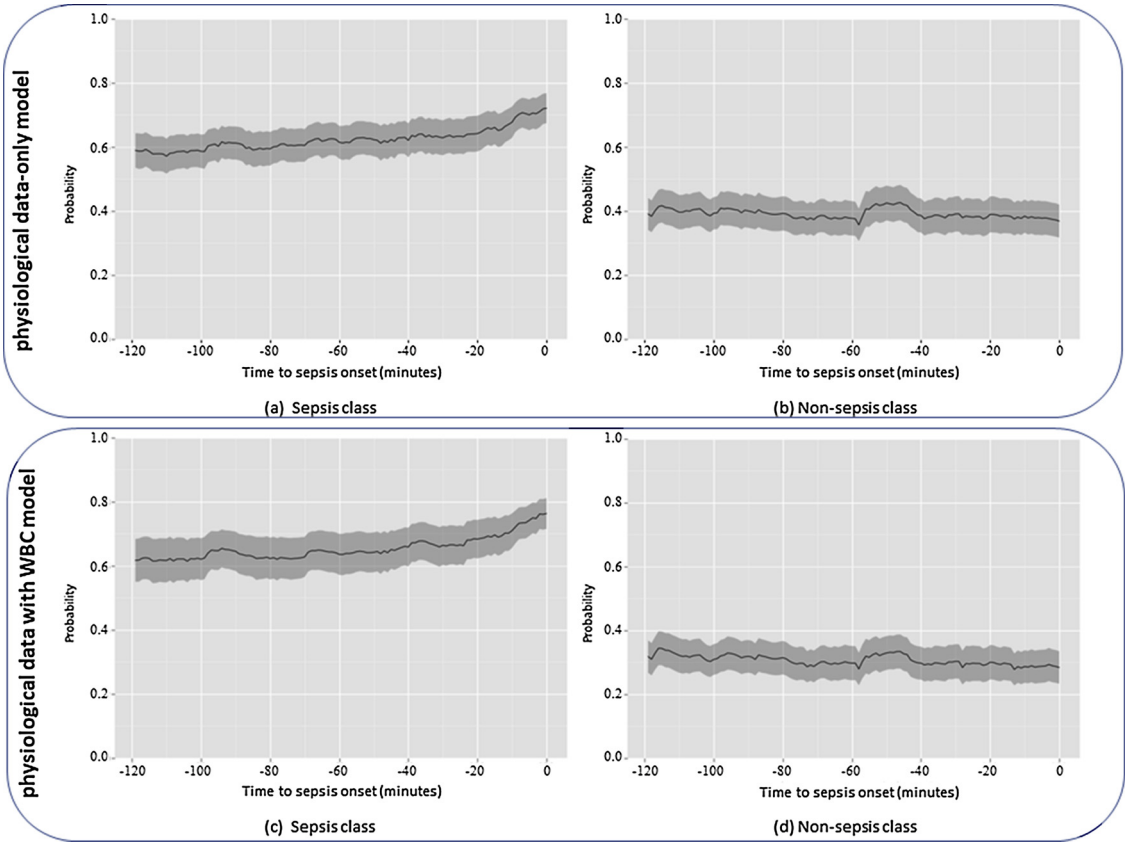

Fig. 6. Mean and 95% confidence interval of the model-estimated probabilities for the test set stratified across the sepsis and non-sepsis subgroups for the 3-hour observational period, for the physiological data-only and physiological data with WBC models.

sepsis patients. This suggests that machine learning features that rely only on physiological data exist that can be predictive of sepsis. However, in our analysis, including WBC increased the specificity of the physiological data-only model. Hence, a model that uses the WBC count along with physiological data can possibly result in fewer false positives (false alarms). Note that in practice, sepsis onset itself may go unnoticed as it relies on continuously calculating SIRS criteria, which is rarely implemented in the highly dynamic clinical environment. Hence, if implemented in practice, the developed automated algorithms can potentially predict sepsis patients much before they are detected at the bedside.

The feature extraction used in this study is novel and inspired by the body's temporal response to infection. Specifically, we extracted features from high-frequency data streams during relatively large time-window, i.e., one hour, to establish a proper baseline. However, we also combined these extracted features with those obtained from shorter time intervals, i.e., the last 30 and 15 min within the time-window, to emphasize the latest physiological changes. The impact and benefit of this feature extraction approach is also evident based on the most significant features contributing to discriminating between sepsis and non-sepsis patients. Our analysis shows that the top physiological features in both models developed using the 3- and 6-hour observational periods are drawn from a mixture of these different interval sizes. Specifically, for the 3-hour observational period, the top physiological features include the maximum respiratory rate from each of the 1, 0.5, and 0.25-hour interval windows, as well as the variance and standard deviation of SBP for the 0.5-hour interval window. Similarly, for the 6-hour observational period, the top physiological features include the maximum respiratory rate from each of the 1, 0.5, and 0.25-hour interval windows, the mean respiratory rate from the 0.25-hour interval window and standard deviation of SBP from the 0.5-hour interval window (see Supplemental Table 4).

Unlike traditional batch processing systems, the presented method was developed to ingest continuous data streams, specifically physiological data streams that arrive from the bed-side. The feature extraction therefore continuously generate new features and the predictive model analyzes those features to generate continuous predictions. As such, this method can reduce the implementation barrier when deployed in an online fashion. Continuous feature extraction and processing can greatly reduce the recognition gap, by allowing computing systems to continuously operate in real-time, thus avoiding the limitations of executions at pre-defined time intervals.

The developed models show promise in the early prediction of sepsis, providing an opportunity for directing early intervention efforts to prevent/treat sepsis cases. However, more study is needed as our work has limitations. In this study, while four hospitals were involved in the data collection, all patients were selected from a single hospital system in the Mid-South USA. Moreover, all patients were admitted to the emergency department, ICU, surgical wards, or other complex care units. Furthermore, the models were developed using the data from patients for whom SIRS criteria could be calculated (see consort diagram illustrated in Fig. 1). We also identified suspicion of infection by referring to the presence of ICD 10 discharge codes for sepsis and severe sepsis. We recognize that it is possible that some patients who developed sepsis during their stay were not coded. Furthermore, we acknowledge that in this work, mainly due to lack of access to all data elements needed for sepsis-3 definition, we rely on an older sepsis definition of sepsis, namely sepsis-2, to identify sepsis onset. In addition, because the exact sepsis onset is not recorded in the EHR, we retrospectively calculate the onset of sepsis using SIRS, so there may also be a disconnect between when SIRS criteria were met and when organ failure, a hallmark of sepsis, occurred. In addition, we use continuous minute-by-minute physiological data streams. Such facilities may not be available at all hospital settings; hence our work only applies to

facilities where streaming data is available. Finally, the models were developed using a retrospective dataset, hence prospective application would reveal significant information about the practical utility and effectiveness.

In conclusion, we illustrate the performance of a state-of-the-art machine learning technique applied to data captured at the bedside with a short collection interval. We develop RF models for various time periods before sepsis onset, in order to discriminate between sepsis and non-sepsis patients. These results demonstrate that salient physiomarkers identified in continuous physiological data streams have the potential to complement decision making at the bedside by non-intrusively predicting patients at the highest risk for developing sepsis.

Contributors' statement page

Dr. Begoli provided informatics specific feedback and critically reviewed the manuscript for important intellectual content.

Dr. Davis conceptualized and designed the study and critically reviewed the manuscript for important intellectual content.

Dr. Kamaleswaran conceptualized and designed the study, developed software to collect and preprocess the data, performed data analysis, and drafted the initial manuscript.

Dr. Khojandi conceptualized and designed the study, supervised data analysis, and participated in the editing of the manuscript.

Dr. Mohammad contributed to the data preprocessing and critically reviewed the manuscript for important intellectual content.

Mr. Van Wyk performed the data analysis and participated in the drafting and editing of the manuscript.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflicts

Dr. Davis received funding from GlaxoSmithKline. The remaining authors have disclosed that they do not have any potential conflicts of interest.

Summary points

What is known:

- Sepsis, characterized by the body's over-reaction to infection, is among the most significant contributors of mortality in intensive care units.
- Predictive algorithms have been shown to significantly advance the time to detection of sepsis, however using data that is captured aperiodically.
- High-frequency data captured by bedside monitors can provide salient 'Physiomarkers' predictive of pathological conditions

What this study adds:

- Features generated from six routinely monitored physiological data streams can predict the onset of sepsis up to 5 h before SIRS, a definition that is actively used across many hospital systems.
- Longer physiological monitoring (6 h vs 3 h) allows for improved predictive times.

Acknowledgments

We would like to acknowledge the efforts of Michael Younker, Brian Williams, Don MacMillan for their work in preparing and providing key data elements that were used in this paper. We would like to thank Dr. David Maslove, Queen's University, Canada for his invaluable feedback throughout the algorithm development.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijmedinf.2018.12.002.

References

[1] F.S. Collins, Varmus H: a new initiative on precision medicine, N. Engl. J. Med. 372 (9) (2015) 793–795.
[2] C.Y. Hung, W.C. Chen, P.T. Lai, et al., Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database, Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE (2017) 3110–3113.
[3] J.L. Jameson, D.L. Longo, Precision medicine—personalized, problematic, and promising, Obstet. Gynecol. Surv. 70 (10) (2015) 612–614.
[4] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, Health Inf. Sci. Syst. 2 (1) (2014) 3.
[5] G. Roesems-Kerremans, Big data in healthcare, J. Healthc. Commun. 1 (2016) 4.
[6] J.S. Upperman, J. Lacroix, M.A. Curley, et al., Specific etiologies associated with the multiple organ dysfunction syndrome in children: part 1, Pediatr. Crit. Care Med. 18 (3_suppl) (2017) S50–7 Mar 1.
[7] Brit Long, Alex Koyfman, Clinical mimics: an emergency medicine–focused review of sepsis mimics, J. Emerg. Med. 52.1 (2017) 34–42.
[8] I. Jawad, I. Lukšić, S.B. Rafnsson, Assessing available information on the burden of sepsis: global estimates of incidence, prevalence and mortality, J. Global Health 2 (1) (2012).
[9] S.P. Shashikumar, M.D. Stanley, I. Sadiq, et al., Early sepsis detection in critical care patients using multiscale blood pressure and heart rate dynamics, J. Electrocardiol. (2017).
[10] L.M. Eerikäinen, J. Vanschoren, M.J. Rooijakkers, et al., Reduction of false arrhythmia alarms using signal selection and machine learning, Physiol. Meas. 37 (8) (2016) 1204.
[11] V.J. Ribas, J.C. Lopez, J.C. Ruiz-Rodriguez, et al., On the use of decision trees for ICU outcome prediction in sepsis patients treated with statins, Computational Intelligence and Data Mining (CIDM), IEEE Symposium on 2011 IEEE, 2011, pp. 37–43.
[12] M. Khalilia, S. Chakraborty, M. Popescu, Predicting disease risks from highly imbalanced data using random forest, BMC Med. Inf. Decis. Making 11 (1) (2011) 51.
[13] G. Litjens, T. Kooi, B.E. Bejnordi, et al., A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[14] D. Kollias, A. Tagaris, A. Stafylopatis, et al., Deep neural architectures for prediction in healthcare, Complex Intell. Syst. (2017) 1–3.
[15] H. Wang, A.C. Roa, A.N. Basavanhally, et al., Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features, J. Med. Imaging 1 (3) (2014) 034003.
[16] A. Khemphila, V. Boonjing, Heart disease classification using neural network and feature selection, 21st International Conference (2011) 406–409 IEEE.
[17] Y. Zhang, Y. Sun, P. Phillips, et al., A multilayer perceptron based smart pathological brain detection system by fractional fourier entropy, J. Med. Syst. 40 (7) (2016) 173.
[18] M.R. Kraft, K.C. Desouza, I. Androwich, Data mining in healthcare information systems: case study of a veterans' administration spinal cord injury population, Proceedings of the 36th Annual Hawaii International Conference (2003) 9-pp. IEEE.
[19] H.J. Kam, H.Y. Kim, Learning representations for the early detection of sepsis with deep neural networks, Comput. Biol. Med. 89 (2017) 248–255.
[20] R. Kamaleswaran, R. Mahajan, O. Akbilgic, A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using varying length single lead electrocardiogram, Physiol. Meas. (2018), https://doi.org/10.1088/1361-6579/aaaa9d.
[21] F. Van Wyk, A. Khojandi, R. Kamaleswaran, et al., How Much data should we collect? A case study in sepsis detection using deep learning, Healthcare Innovations and Point of Care Technologies (HI-POCT) (2017) 109–112.
[22] S. Nemati, A. Holder, F. Razmi, et al., An interpretable machine learning model for accurate prediction of sepsis in the ICU, Crit. Care Med. 46 (4) (2017) 547–553.
[23] K.E. Henry, D.N. Hager, P.J. Pronovost, et al., A targeted real-time early warning score (TREWScore) for septic shock, Sci. Transl. Med. 7 (299) (2015) 299ra122.
[24] M.P. Griffin, T.M. O'Shea, E.A. Bissonette, et al., Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness, Pediatr. Res. 53 (6) (2003) 920.
[25] S. Ahmad, T. Ramsay, L. Huebsch, et al., Continuous multi-parameter heart rate variability analysis heralds onset of sepsis in adults, PloS one 4 (8) (2009) e6642.
[26] R. Kamaleswaran, et al., Applying artificial intelligence to identify physiomarkers predicting severe sepsis in the PICU, Pediatr. Crit. Care Med. 19.10 (2018) e495–e503.
[27] Cerner Corporation. CareAware iBus. Available https://www.cerner.com/pages/careaware. Accessed March 13, 2018.
[28] Centers for Disease Control and Prevention, International Classification of Diseases,

Tenth Revision, Clinical Modification (ICD-10-CM), Atlanta, Georgia, USA. Available on (2018) https://www.cdc.gov/nchs/icd/icd10cm.htm.
[29] M. Sartelli, Y. Kluger, L. Ansaloni, et al., Raising concerns about the sepsis-3 definitions, World J. Emerg. Surgery WJES 13 (2018) 6, https://doi.org/10.1186/s13017-018-0165-6.
[30] M. Barati, F. Alinejad, M.A. Bahar, M.S. Tabrisi, A.R. Shamshiri, H. Karimi, Comparison of WBC, ESR, CRP and PCT serum levels in septic and non-septic burn cases, Burns 34 (6) (2008) 770–774.
[31] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[32] A.D. Gordon, L. Breiman, J.H. Friedman, et al., Classification and regression trees, Biometrics 40 (3) (1984) 874.
[33] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[34] T.K. Ho, Random decision forests (PDF), Proceedings of the 3rd International Conference on Document Analysis and Recognition (1995) 278–282 14–16.
[35] M.H. Hassoun, Fundamentals of Artificial Neural Networks, MIT press, 1995.
[36] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[37] Apache Spark: Lightning-Fast Cluster Computing, (2016) Available on https://spark.apache.org/.
[38] T.E. Oliphant, Python for scientific computing, Comput. Sci. Eng. 9 (3) (2007) 10–20.
[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al., Scikit-learn: machine learning in python, J. Mach. Learn. 12 (2011) 2825–2830, https://doi.org/10.1007/s13398-014-0173-7.2.
[40] F. Chollet, Keras: The Python Deep Learning Library, Astrophysics Source Code Library, 2018.
[41] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, Tensorflow: a system for large-scale machine learning, OSDI 16 (2016, November) 265–283.