Prediction of Septic Shock Onset in ICU by Instantaneous Monitoring of Vital Signs*

Maximiliano Mollura[1**], Stefano Romano[1**], Giulio Mantoan[1],
Li-wei Lehman[2] *Senior Member, IEEE* and Riccardo Barbieri[1] *Senior Member, IEEE*

***Abstract*— Septic Shock is a critical pathological state that affects patients entering the intensive care unit (ICU). Many studies have been directed to characterize and predict the onset of the septic shock, both in ICU and in the Emergency Department employing data extracted from the Electronic Health Records. Recently, machine learning algorithms have been successfully employed to help characterize septic shock in a more objective and automatic fashion. Only a few of these studies employ information contained in the continuously recorded vital signs such as electrocardiogram and arterial blood pressure. In particular, we have devised a novel feature estimation procedure able to consider instantaneous dynamics related to cardiovascular control. This work aims at developing a short-term prediction algorithm for identifying patients experiencing septic shock among a population of 100 septic patients extracted from the MIMIC-III clinical and waveform database. Among all the results obtained from several trained machine learning models, the best performance reached an AUC on the test set equal to 0.93 (Accuracy=0.85, Sensitivity=0.89 and Specificity=0.82).**

***Clinical Relevance*— The study paves the way for instantaneous monitoring of ICU patients focusing on predicting septic shock onset, a critical problem that needs to be addressed in this clinical setting.**

I. INTRODUCTION

Patients admitted to the Intensive Care Unit (ICU) represent the most critical population in the hospital, characterized by a wide spectrum of diseases and conditions, with the only common denominator of life threatening and continuous monitoring. This continuous monitoring provides a large amount of data, stored in the electronic health records (EHR), and continuously recorded vital signs, such as electrocardiogram (ECG) and arterial blood pressure (ABP) waveforms

According to the third international consensus held in 2016 [1], Sepsis is one of the illnesses which has the highest impact to critical ill patients in terms of mortality and, within septic patients, the most critical condition is represented by the septic shock, a state of acute circulatory failure associated with infection and with higher mortality [2]. Therefore, patients who develop septic shock are particularly critical and must be treated with great attention to avoid a deadly outcome. Several studies have demonstrated that morbidity, mortality and length of ICU stay are decreased when septic shock is identified and treated early [3][4]. Therefore, prediction of the onset of Septic Shock has been investigated by several studies in last years.

A first approach was proposed by Shavdia in 2007 [5], with a multivariate logistic regression model, using different prediction times before hypotension onset, considered here as onset of the shock. That study considered features coming from laboratory measures and waveforms, leading to AUCs of about 0.94 and an accuracy of 0.86. Other research, [6][7], analyzed patients from the emergency department, trying to identify septic shock patients before their admission in ICU, so predicting with a very large advance the onset.

The TREWScore, proposed by Henry et al. [8], consists of a real-time early warning score that identifies patients at high risk of developing septic shock. The score considers features from laboratory measures and clinical values of heart rate and blood pressure, so routinely available in emergency department, labeling the patient at risk when the score overcomes the estimated threshold. The score reached performance of about 0.83 of AUC in validation set, with patient identified the day before shock onset.

Another approach, proposed by Darwiche et al. [9], focuses on application of machine learning techniques, building up a Cox Enhanced Random Forest (CERF) algorithm. Starting features considered were only laboratory measures and vitals such as heart rate, systolic and diastolic blood pressure, respiration rate and temperature. The prediction is made 20 hours before the real onset, and achieves good performances, with a sensitivity of almost 89% and a specificity of 97%. The main limitations of this study can be found in the exclusion of patients who received extensive treatments and patients who had septic shock onset in first 5 hours of ICU stay, due to lack of recorded data.

This work aims at developing algorithms for the short-term prediction of the septic shock onset using only features extracted from the continuous records of ECG and ABP, employing machine learning classifiers.

II. MATERIALS AND METHODS

A. Cohort Selection

The data used in this study are extracted from the MIMIC-III Waveform and Clinical Databases, publicly available on Physionet, containing data recorded in the Beth Israel Deaconess Medical Center ICU from 2001-2012 with two different systems, Carevue (CV), whose data range form

*This work was partially supported by Progetto Roberto Rocca, MIT-Italy Program

**These authors equally contributed to the development of this work

[1] Stefano Romano, Maximiliano Mollura, Giulio Mantoan and Riccardo Barbieri are from the Department of Electronic, Information and Biomedical Engineering, Politecnico di Milano. maximiliano.mollura@polimi.it

[2] Li-wei Lehman is from the Laboratory for Computational Physiology, Massachusetts Institute of Technology

TABLE I
DEMOGRAPHICS AND FRACTIONS OF CO-MORBIDITIES OF THE CONSIDERED COHORT (S=SEPSIS, SS=SEPTIC SHOCK GROUPS).

Demographics					
	Gender (F)	Age (yr)	LOS (days)	Hosp.	28 days
S	0.5	56(43-62)	2.1(1.4-4.3)	0.07	0.07
SS	0.5	55(47-63)	6.7(2.6-14.4)	0.22	0.22

Co-morbidities					
	CHF	Diab.	RF	LD	CGPT
S	0.04	0.22	0.02	0.16	0.11
SS	0.04	0.18	0.02	0.22	0.22

2001-2008 and Metavision (MV) systems from 2008-2012 [11].

The availability of the clinical information related to the bedside recorded vital signs allowed to define the septic shock onset according to the "Third Consensus on Sepsis and Septic Shock" [1] as the contemporaneous administration of vasopressors and a measure of Serum Lactate level>2mmol/L for septic patients. Consequently, starting from a cohort of septic subject already matched with the waveform database (n=2068) the following criteria were applied:

- Only data registered with MV system are considered because of the more accurate registration of administrations.
- Presence of both electrocardiogram (ECG), either I, II or "V" lead, and arterial blood pressure (ABP) waveforms.
- Availability of 1-hour recording before the septic shock onset, defined as the administration of vasopressors and contemporary Serum Lactate $>$ 2 mmol/L (for patients classified as *shock*).
- Availability of 1-hour recording elsewhere in the first 24 hours of ICU stay for patients classified as *NoShock*.
- Age higher than 18 years old at the admission in ICU

The final cohort consisted on 100 subjects, 45 *Shock* and 55 *NoShock* patients. Septic patients, were identified thanks to the clinical records according to the following criteria, increase in Sequential Organ Failure (SOFA) greater than 2 - or quick SOFA (qSOFA) greater than 2 - and contemporaneous administration of antibiotic therapy [12], as supection of infections marker, as reported in the third definition of Sepsis [1]

The main characteristics of the considered cohort are shown in table I, where it can be noted that both septic (S) and septic shock (SS) populations had 50% of female patients (Gender) and a similar distribution of ages (Age); of note, septic shock patients show a higher length of stay (LOS), hospital and 28-days mortalities. In the table are also shown co-morbidities of the included patients as congestive heart failure (CHF), Diabetes (Diab), Renal Failure (RF), Liver Disease (LD) and the presence of coagulopathy (CGPT).

B. Feature Extraction

For each patient, ECG and ABP waveforms were annotated extracting the time of occurrence of R-peaks in the ECG with the Pan-Tompkins algorithm [13] and the occurrence and absolute value of systolic and diastolic phases from ABP waveform [14]; the pressure onset was then defined as the point at maximum derivative between the R-peak and the systolic phase. The annotations were then synchronized to identify a correspondence between the R-peak and the fiducial points of the blood pressure.

From the extracted annotations it was possible to build the following time series: the tachogram as the series of time intervals between successive R-peaks (RR), the systogram as the series of successive systolic (SAP) values and the diastogram as the series of diastolic values (DAP) following the systolic ones.

Moreover, the pulse arterial pressure (PPress), mean arterial pressure (MAP) and pulse arrival times (PAT) time series were extracted as follows:

$$PPress(i) = SAP(i) - DAP(i) \tag{1}$$

$$MAP(i) = \frac{SAP(i) + 2DAP(i)}{3} \tag{2}$$

$$PAT(i) = O_t(i) - R(i) \tag{3}$$

where O_t is the series of pressure onset time occurrences.

From the continuously recorded vital signs of each subject were extracted features able to describe the state of the patient. In particular, from the RR series we extracted measures about the HRV, both linear (AVNN, SDNN, SDANN, NN20, NN50, RMSSD, TRI), non-linear (SD1, SD2, SDratio, Sample Entropy, Lyapunov exponents, long-term coefficient of DFA) and spectral (LF, HF, LF/HF, LFn, HFn). Spectral components were extracted also from systolic and diastolic blood pressure series.

Additionally, R-peaks occurrences and systolic values were used with a bivariate Point-Process modeling [15] to estimate the average RR interval ($\mu_{RR}(t)$) with a 200Hz time resolution and the coefficients of the considered model (a_i, b_j) were used to compute time varying spectral measures. From these time varying spectral measures were computed their first, second, third and fourth statistical moments and the slope of their regression lines.

The general expression of the autoregressive bivariate model is:

$$\mu_{RR}(t) = a_0 + \sum_{i=1}^{p} a_i RR(t-i) + \sum_{j=1}^{q} b_j SAP(t-j) \tag{4}$$

where its parameters are estimated by maximizing the log-likelihood of a history dependent inverse gaussian probability, whose probability density function can be expressed as follows:

$$p(t) = \left(\frac{\theta}{2\pi(t-u_j)^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\theta(t-u_j-\mu_{RR}(t))}{2\mu_{RR}(t)^2(t-u_j)}\right) \tag{5}$$

for each $t > u_j$, where $(t - u_j)$ represents the time from the last occurred R-peak and $\theta > 0$ a shape parameter. The order of the autoregressive bivariate model was empirically set to p=q=13 for all patients, by evaluating the performances in terms of KS-distance and autocorrelation function, identifying a trade-off between these two measures of goodness-of-fit.

From all the extracted time series were computed their first, second, third and fourth statistical moments and the slope of their regression lines in order to extract information about time evolution and possible trends.

The study was set up in order to predict septic shock 15 minutes in advance with respect to the real shock onset. Therefore, we extracted features from a data window of 45 minutes with 15 minutes of lead time before the shock.

C. Model Selection and Prediction

Due to the large number of extracted features, a feature selection step was required before the training of any algorithm. After splitting training and test set with a stratified 80%-20% partition, we used a forward selection on the features of the training set to extract the most important features that can describe the dataset.

Then, the features were used to train Logistic regression (LR), Trees, Support Vector Machines (SVM), k-Nearest Neighbors (kNN) and Ensemble Tree (E.TREE) classifiers in order to determine whether a patient would develop a shock in the next 15 minutes or not. The parameters of the model were estimated with a 10-fold cross-validation (CV). The hyperparameters of each model were trained using a Bayesian optimization rule Also in this case a 10-fold CV on the training set was performed.

Trained models were tested on the test set and performances were evaluated using the Area under the receiving operating characteristic curve (AUC), Sensitivity (SE), Specificity (SP), Accuracy (ACC), F1 score (F1), Positive Predictive Value (PPV) and Negative Predictive Value (NPV).All analyses were performed with MATLAB 2018b.

III. Results

For all the considered data, the point process model achieved satisfactory performances in model goodness-of-fit.

After the feature selection step, the extracted features were: normalized low frequency and high frequency components of the systolic series (SAP_{LFn} and SAP_{HFn}), the ratio between low frequency and high frequency components of the SAP ($SAP_{LF/HF}$), the mean value of the pulse pressure ($PPress_{avg}$) and from the set of features extracted within the point process modeling framework, the skewness of the high frequency component of the RR time varying spectrum ($RR_{PP,HFskew}$) and the slope of the ratio between low and high frequency components extracted from the coefficients of the point process feedback branch ($SAP_{PP,LF/HFslope}$).

Three out of five classifiers achieved high AUC (>=0.90), see Fig. 1, and SE (>=0.89), except for tree based methods. LR achieved the best performance in terms of AUC, with 0.93, with good results also in terms of PPV (0.80) and NPV (0.90) and achieving an accuracy of 0.85.

The kNN prediction model showed the best performances in terms of F1 score, NPV and SE, obtained correctly identifying all shock patients with 3 out of 11 misclassified controls. Overall results are shown in table II.

TABLE II

Results obtained in Sepsis Identification on the test set with different classification algorithms.

Identification Results					
	LR	SVM	kNN	TREE	E.TREE
AUC	**0.93**	0.91	0.90	0.72	0.88
F1	0.84	0.82	**0.86**	0.70	0.80
PPV	**0.80**	0.69	0.75	0.64	0.73
NPV	0.90	**1**	**1**	0.78	0.89
SE	0.89	**1**	**1**	0.78	0.89
SP	**0.82**	0.64	0.73	0.64	0.73
ACC	**0.85**	0.80	**0.85**	0.70	0.80

IV. Discussion and Conclusions

Results obtained in this study suggest that 45-minutes recording of commonly used vital signs to monitor patients in the ICU are able to predict whether a patient will develop a septic shock 15-minutes in advance of its onset. In addition to this result, the availability of instantaneous features allows for a further important characterization of the physiopathological mechanisms leading to the shock.

To this extent, the features employed to train the different machine learning algorithms cover different aspects of the cardiovascular control loop, including its dynamical characteristics. In particular, the modeling approach allows the extraction of indexes related to the temporal dynamics of standard spectral measures

Final models include estimates from the RR interval derived point process indices, as well as measures extracted from systolic and pulse arterial pressure time series. The included features suggest that blood pressure plays a key role in identifying shock patients both in terms of absolute values, as evidenced by the $PPress_{avg}$ and in terms of the autonomic control. Indeed, low frequency oscillations computed from the systolic time series might be related to autonomic control and vasomotor tone regulation, as suggested by [16], which might be influenced when evolving in a shock condition. In addition, as observed by [17], high frequency components extracted from the blood pressure time series would carry additional information about cardiovascular changes elicited by respiratory activity.

Of note, $SAP_{LF/HF}$ and $SAP_{PP,LF/HFslope}$ are among the variables that have great importance in predicting the shock, pointing at a possible effect of the overall dynamic range of the autonomic tone including vascular and respiratory influences.

The importance of using a sound statistical model for heartbeat dynamics is supported by the importance of features related to the spectral information obtained using the point process modeling approach, as related to the dynamic evolution of autonomic modulation in this critical population. In particular, $RR_{PP,HFskew}$ highlights the changes in the distribution of high frequency oscillations (i.e. vagal modulation) while approaching the shock. On the other hand, the $SAP_{PP,LF/HFslope}$ conveys information about trends in the influences of the feedback branch of the loop on the heart, which also include the baroreflex, and their temporal

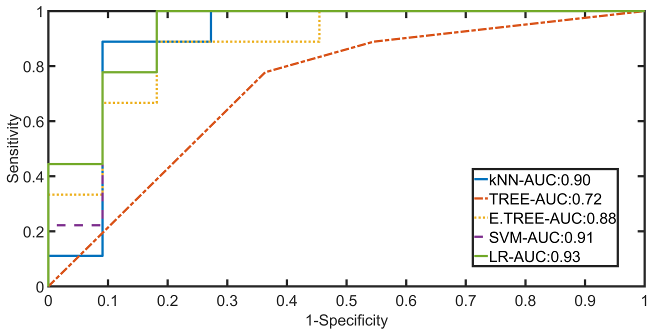

Fig. 1. ROCs obtained on the test set with kNN, Tree, Ensemble Tree, SVM and Logistic Regression algorithms.

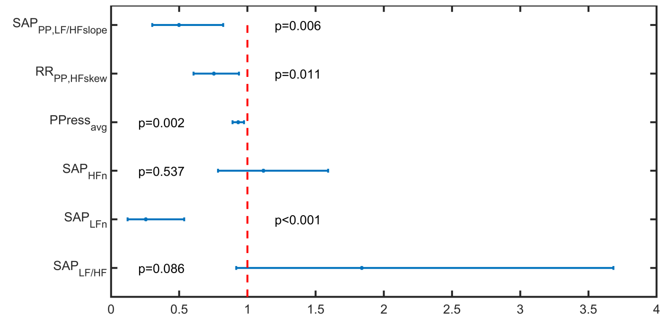

Fig. 2. Odds ratios and 95% confidence intervals of features included in the Logistic Regression classifier.

evolution towards the onset.

Thanks to the logistic regression model it is also possible to extract odds ratios, shown in Fig. 2, and p-values of each variable used to fit the model. As a result, $SAP_{PP,LF/HFslope}$, SAP_{LFn}, $RR_{PP,HFskew}$ and $PPress_{avg}$ were revealed to be significant in the multivariate model ($p < 0.05$), strengthening the importance of these features in predicting shock onset.

Our overall prediction results are numerically comparable with other studies (AUC=0.93 vs a maximum of 0.94 [5]).

Of note, our prediction window is scaled to the amount of information provided by high temporal resolution waveforms. We have a shorter estimation margin but a more accurate prediction in time. Other successful techniques, such as [8],[7], mostly use scores that are not linked to a precise time of the shock event and extract information only from EHRs. In addition our procedure does not link the onset definition to the hypotensive events, which might not be observed if a vasopressor therapy is promptly initiated, thus allowing for a higher number of subject that can be included in the study. Our procedure also includes patients undergoing extensive treatments. The other main innovation of our study is indeed represented by the fact that only waveform recordings are considered, with a definition of shock onset based on contemporaneous administration of vasopressor and Lactate>2. Our results need to be further validated, increasing the prediction window and the cohort size. The ability of our algorithm to discriminate septic shock from the other shock states, as cardiogenic or hypovolemic shocks will be also investigated in future studies.

References

[1] M. Singer, C. S. Deutschmann, C. W. Seymour et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):801-810. doi:10.1001/jama.2016.0287

[2] Shankar-Hari M, Phillips GS, Levy ML, et al. Developing a New Definition and Assessing New Clinical Criteria for Septic Shock: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):775–787. doi:10.1001/jama.2016.0289

[3] H. B. Nguyen, S. W. Corbett, R. Steele et al. Implementation of a bundle of quality indicators for the early management of severe sepsis and septic shock is associated with decreased mortality. Critical Care Medicine. 2007;35(4):1105-1112. doi:10.1097/01.CCM.0000259463.33848.3D

[4] V. Coba, M. Whitmill, R. Mooney et al. Resuscitation bundle compliance in severe sepsis and septic shock: Improves survival, is better late than never. Journal of Intensive Care Medicine. 2011;26(5):304-313. doi:10.1177/0885066610392499

[5] D. Shavdia. Septic shock: Providing early warnings through multivariate logistic regression models. 2007; MD thesis, Massachusetts Institute of Technology

[6] S. W. Thiel, J. M. Rosini, W. Shannon et al. Early prediction of septic shock in hospitalized patients. Journal of hospital medicine: an official publication of the Society of Hospital Medicine. 2010;5(1):19-25. doi:10.1002/jhm.530

[7] J. Kim, H. Chang, D. Kim, D.H. Jang, I. Park, K. Kim. Machine learning for prediction of septic shock at initial triage in emergency department. Journal of Critical Care, 2020. 55:163-170. https://doi.org/10.1016/j.jcrc.2019.09.024.

[8] K. E. Henry, D. N. Hager, P. J. Pronovost eh al. A targeted real-time early warning score (trewscore) for septic shock. Science Translational Medicine. 2015;7(299). doi:10.1126/scitranslmed.aab3719

[9] A. Darwiche, S. Mukherjee. Machine learning methods for septic shock prediction. In Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality, AIVR 2018; (104-110). doi: 10.1145/3293663.3293673

[10] Chen, Z., Purdon, P. L., Harrell, G., Pierce, E. T., Walsh, J., Brown, E. N., Barbieri, R. Dynamic assessment of baroreflex control of heart rate during induction of propofol anesthesia using a point process method. Annals of biomedical engineering, 2011; 39(1), 260–276. doi:10.1007/s10439-010-0179-z

[11] A. E. W. Johnson, T. J. Pollard, L. Shen et al. Mimic-iii, a freely accessible critical care database. Scientific Data. 2016; 3(1):16-35. doi: 10.1038/sdata.2016.35

[12] Johnson AEW, Aboab J, Raffa JD, et al. A Comparative Analysis of Sepsis Identification Methods in an Electronic Database. Crit Care Med. 2018;46(4):494–499. doi:10.1097/CCM.0000000000002965

[13] J. Pan, W.J. Tompkins. A real-time qrs detection algorithm. IEEE Trans. Biomed. Eng. 1985; 32(3):230-236. doi: 10.1109/TBME.1985.325532

[14] B.N. Li, M. C. Dong, M.I. Vai. On an automatic delineator for arterial blood pressure waveforms. Biomedical Signal Processing and Control. 2010; 5(1):76-81. doi:10.1016/j.bspc.2009.06.002

[15] Z. P. Chen, P. L. Harrel, P. Graceand et. al. Dynamic assessment of baroreflex control of heart rate during induction of propofol anesthesia using a point process method. Annals of biomedical engineering. 2011; 39(1):260-276. doi:10.1007/s10439-010-0179-z

[16] G. Parati , J. P. Saul , M. Di Rienzo , and G. Mancia.Spectral Analysis of Blood Pressure and Heart Rate Variability in Evaluating Cardiovascular Regulation. Hypertension. 1995; 6:1276-1286. doi:10.1161/01.HYP.25.6.1276

[17] R. Poletto, A. M. Janczak, R. M. Marchant-Forde, J. N. Marchant-Forde, D. L. Matthews, C. A. Dowell, D. F. Hogan, L. J. Freeman, D. C. Lay. Identification of low and high frequency ranges for heart rate variability and blood pressure variability analyses using pharmacological autonomic blockade with atropine and propranolol in swine. Physiology & Behavior, 2011; 103(2):188-196. https://doi.org/10.1016/j.physbeh.2011.01.019.