

A Data-Driven Approach to Predicting Septic Shock in the Intensive Care Unit

Christopher R Yee, Niven R Narain, Viatcheslav R Akmaev
and Vijetha Vemulapalli

Berg LLC, Framingham, MA, USA.

Biomedical Informatics Insights
Volume 11: 1–9
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1178222619885147



ABSTRACT: Early diagnosis of sepsis and septic shock has been unambiguously linked to lower mortality and better patient outcomes. Despite this, there is a strong unmet need for a reliable clinical tool that can be used for large-scale automated screening to identify high-risk patients. We addressed the following questions: Can a novel algorithm to identify patients at high risk of septic shock 24 hours before diagnosis be discovered using available clinical data? What are performance characteristics of this predictive algorithm? Can current metrics for evaluation of sepsis be improved using novel algorithm? Publicly available data from the intensive care unit setting was used to build septic shock and control patient cohorts. Using Bayesian networks, causal relationships between diagnosis groups, procedure groups, laboratory results, and demographic data were inferred. Predictive model for septic shock 24 hours prior to digital diagnosis was built based on inferred causal networks. Sepsis risk scores were augmented by de novo inferred model and performance was evaluated. A novel predictive model to identify high-risk patients 24 hours ahead of time, with area under curve of 0.81, negative predictive value of 0.87, and a positive predictive value as high as 0.65 was built. The specificity of quick sequential organ failure assessment, systemic inflammatory response syndrome, and modified early warning score was improved when augmented with the novel model, whereas no improvements were made to the sequential organ failure assessment score. We used a data-driven, expert knowledge agnostic method to build a screening algorithm for early detection of septic shock. The model demonstrates strong performance in the data set used and provides a basis for expanding this work toward building an algorithm that is used to screen patients based on electronic medical record data in real time.

KEYWORDS: Artificial intelligence, machine learning, Bayesian networks, electronic health records, sepsis, predictive algorithm, hospital-acquired infections, intensive care unit

RECEIVED: September 19, 2019. **ACCEPTED:** September 23, 2019.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by Berg LLC.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The authors are employees of Berg LLC. N.R.N. is also a director and shareholder of Berg LLC.

CORRESPONDING AUTHOR: Niven R Narain, Berg LLC, 500 Old Connecticut Path, Building B, Framingham, MA 01701, USA. Email: niven.narain@berghhealth.com

Introduction

Sepsis is a serious condition that arises when the body's response to infection causes injury to its own tissues and organs and is a great cause of concern among hospitalized patients.¹ In sepsis, a systemic inflammatory response syndrome (SIRS) is caused by the infection. Sepsis can progress into severe sepsis accompanied by remote organ dysfunction. Septic shock is the most severe form of sepsis that involves presence of arterial hypotension and is associated with significantly worse outcomes.² The incidence rate of sepsis was found to be around 6% in 2014 with about 15% of those patients dying as well as 6.2% of those being admitted to hospice care.³ In addition to the high incidence of sepsis in hospitals, high mortality rates, it is among the most expensive illnesses to treat in the United States. It costs a median of around US\$32 421 per patient for the entire hospital stay as well as US\$27 461 for the intensive care unit (ICU) costs.⁴

As of now, there is no gold standard for diagnosis and prediction of sepsis and septic shock. Currently, septic shock is predominantly diagnosed based on a combination of lab tests and clinical features such as fever, low blood pressure, and difficulty breathing, although some sepsis alert systems exist.^{5–8} Early diagnosis of this condition is complicated by nonspecific clinical signs and symptoms and the fact that not all infections lead to sepsis and to progression into septic shock.¹ The current standard of care for septic shock includes administration of

antibiotics, antifungal drugs, regulating blood volume, and ensuring sufficient tissue perfusion. Surgical source control is a measure that is used less often and is recommended at the earliest possible time to obtain the best outcome.⁹ Early detection of sepsis and early administration of antibiotic treatment have been known to be the strongest modulators of outcomes in patients with sepsis.¹⁰ Although highly desirable, early diagnosis is also more challenging to accomplish in a clinical setting as discussed earlier. Given the nonspecific nature of early symptoms, it is impractical to closely monitor all patients in the ICU. A screening algorithm that identifies patients at high risk of septic shock would enable both higher rates of early diagnosis of septic shock and better utilization of often scarce clinical resources. Such an algorithm would limit close monitoring to a much smaller patient population, making it both practicable and invaluable in the ICU setting.

Many prior efforts have been made by other research groups to develop algorithms to predict sepsis using electronic medical record data from MIMIC or other patient databases^{11–17}; commercial early warning systems are also available.¹⁸ It remains challenging to benchmark and compare performance of different methods due to variation in methods used to identify patient cohorts from electronic medical record data and in clinical definitions of sepsis. The research presented in this article was performed with the objective of using data-driven methods that are unbiased by expert knowledge to identify a



predictive algorithm for predicting progression into septic shock.

The aim of this work is to use publicly available ICU data to develop a screening algorithm for early detection of septic shock. The algorithm built in this work identifies high-risk patients 24 hours prior to a digital diagnosis of septic shock. A diagnosis that was made by applying certain criterion to electronic medical record data is referred to as a “digital diagnosis” within the scope of this article. A data-driven approach to developing the risk stratification algorithm was adopted through use of advanced machine learning and artificial intelligence-based techniques. Data-driven approaches are unbiased by current knowledge and provide an alternative to expert knowledge-based, hypothesis-driven approaches. An anonymized clinical database containing electronic medical of close to 50 000 patients in the ICU setting from the Beth Israel Deaconess Medical Center was used for this analysis.^{19,20} Performance characteristics of the novel risk stratification algorithm are described in this article. We also attempted to augment currently used sepsis scores with the newly identified risk factors and aim to improve performance of the scoring systems.

Methods

Data

The MIMIC-III Clinical Database^{19,20} contains health records on ICU admissions from 2001 to 2012 at the Beth Israel Deaconess Medical Center. This data set is publicly available²¹ and contains a wide range of medical data including laboratory test results, vital signs, diagnosis, and procedure codes. The data were imported from CareVue clinical information system²² for admissions between 2001 and 2008 and MetaVision (provided by iMDSoft) between 2008 and 2012. The local copy of the database was hosted on a Hadoop²³ cluster and was queried using Hive.²⁴ R language²⁵ was used for the data processing and statistical analysis.

Criterion for identification of patients who have progressed into sepsis shock

A working definition for digital diagnosis of septic shock is described in this section. The definition provided here is used as the “gold standard” for (1) building a novel predictive algorithm for septic shock, (2) evaluating performance of a novel predictive algorithm for septic shock, and (3) evaluating performance of widely used scoring systems for sepsis. The criterion for identifying patients in septic shock from electronic medical record data was based on an earlier work by Kadri et al²⁶ The algorithm proposed by Kadri et al was adopted to enable use with the MIMIC-III database.

Figure 1 describes the timelines and the clinical criterion used to identify progression to septic shock in a patient during their stay in the ICU. A patient is considered to have gone into septic shock on day 0 if the following conditions are met: (1)

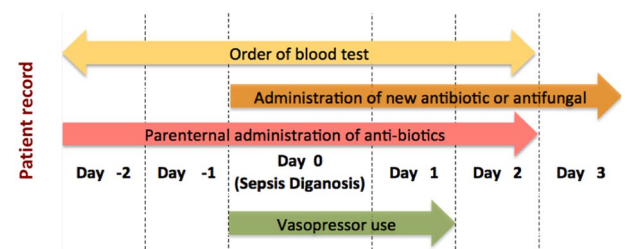


Figure 1. Occurrence of septic shock in each patient is identified based on 2 conditions: (1) Presence of an indicator of shock (vasopressor use) and (2) at least one indicator of presumed infection: blood culture test, administration of antibiotics or antifungals. Day 0 is considered to be the day of onset of septic shock. Timelines for identification of conditions are indicated in the figure.

Relevant vasopressor use occurs on day 0 and day 1. Vasopressor drugs that were considered are norepinephrine, epinephrine, vasopressin, dopamine, and phenylephrine. Vasopressors were identified from the input events and chart events tables in the MIMIC-III database. (2) Relevant blood culture orders were made from day -2 to day 2 relative to the onset of septic shock. Blood culture orders were identified from the microbiology events table from the database. (3) Administration of a new antibiotic or antifungal from day 0 to day 3 is considered relevant to identification of onset of septic shock. Similarly, parenteral administration of antibiotic between day -2 and day 2 is also considered relevant. Drug administration was identified from the input events table based on National Drug Codes (NDCs). National Drug Codes from the MIMIC-III database were mapped to the complete NDC directory as listed by the Food and Drug Administration.²⁷ The drug class categories as defined in the NDC directory were used to identify 2704 antibacterial and 560 antifungal drugs. Based on route of administration information from the NDC directory, 1252 antibacterial and 226 antifungal drugs were identified to characterize parenteral administration of antimicrobials. Details corresponding to identification of various features listed here are available in the supplementary material.

Patient cohort selection

A cohort of “septic shock patients” or cases was built by applying the algorithm for digital diagnosis of septic shock as described previously. A control cohort was identified by randomly sampling from all admissions in the database not including patients who have been identified as progressing into septic shock. No matching at the individual patient level was performed—a simple random sampling approach was taken. For each patient in the control cohort, a time point during their ICU stay was chosen for notional diagnosis of septic shock. The time points for notional diagnosis of septic shock in control patient population was chosen as follows: (1) the number of hours of hospital stay prior to diagnosis of septic shock was identified for each patient in the septic shock cohort and (2) time point of notional septic shock for each patient in the

control cohort was sampled from the time points calculated from the septic shock cohort.

For each patient in the control cohort, a time point for notional diagnosis of sepsis was chosen as described above. A time point for notional diagnosis of septic shock was chosen to temporally align patients from the septic shock and control cohorts. Patients in the control cohort do not have a diagnosis of septic shock at the time of notional septic shock, however, relevant lab and diagnosis code data for the control cohort patients was determined based on the time of notional septic shock.

Data processing for construction of Bayesian networks

For patients in the septic shock and control cohort, data relevant to lab events, specifically the lab test item ID and the results from manual interpretation, were extracted from the MIMIC-III database for all visits. The extracted lab data were inherently discrete with the following levels: normal, abnormal, delta (a large or sudden change of a lab result from the previous test result), and missing data. If a lab measurement was not made for the time window under consideration, the lab measurement was interpreted as “not measured.” The *International Classification of Diseases, Ninth Revision* (ICD-9) diagnosis and procedure codes for both cohorts were extracted from the MIMIC-III database. The large number of ICD-9 codes were reduced using the Clinical Classification Software²⁸ (CCS) algorithms from the Agency for Healthcare Research and Quality. The CCS data set allows for group similar diagnostic and procedure codes to increase the density of the data and improve interpretability of the codes.

A data set representing snapshots of the medical record of the patients 24 hours prior to the diagnosis of septic shock was created. For lab data, to increase the completeness of the patient profiles, all lab result data between 24 and 30 hours prior to diagnosis of septic shock was considered. For diagnosis and procedure codes, if the code was observed any time prior to 24 hours preceding septic shock diagnosis, it was assigned as “present.” If the diagnosis or procedure code was not observed any time prior to the time point, it was assigned as “absent.”

Model building using bAIcIs

An integrated data set containing lab data, diagnosis, procedure, and observation of septic shock was created containing all patients in both the septic shock and control. A Bayesian network containing inferred cause-and-effect relationships between variables in the data set was built using bAIcIs. The method has been described in previous studies.^{29,30} Demographic information was considered to be fixed information regarding the patients. Therefore, in the causal network, no other variables were permitted to drive changes in the demographic characteristics of the patients. Due to temporal nature of the septic shock

diagnosis in relation to other data elements, the diagnosis of septic shock was not permitted to causally drive changes in other variables in the data set. An ensemble of 500 causal networks were built and the results were summarized as a single representative network. The edges (edges link variables/nodes in network) were weighted based on the frequency of observation in the ensemble of networks. The summarized causal network was then filtered to retain only interactions that were observed with a frequency of 0.4 or greater. A subnetwork was selected from the summarized network to include risk factors that drive progression into septic shock. A regression model was built using factors from the subnetwork around the outcome of septic shock. Multiple thresholds were selected to obtain the most desirable model performance in a clinical application. Factors included in the regression model were manually selected from the subnetwork based on their power to predict septic shock.

Calculation of other sepsis scores

Sequential organ failure assessment (SOFA), quick SOFA (qSOFA), modified early warning score (MEWS), and SIRS were calculated. Based on the scores calculated, patients were classified as “high risk of septic shock,” “low risk of septic shock,” or “NA” (score could not be calculated). The SOFA scores³¹ were calculated based on published guidelines using data from MIMIC-III. Data from the chart events, lab events, medication, and diagnosis tables were used. Relevant item IDs corresponding to clinical measure of interest such as coagulation were identified through text searches and manual matching. Six categories (Glasgow Coma Scale [GCS], liver function, coagulation, renal function, respiratory function, and cardiovascular function) were evaluated and the total SOFA score was calculated. Changes in SOFA score could not be estimated from the data due to lack of resolution in the data set. Alternatively, a threshold of 4 was chosen to assign patients to the “high risk of septic shock” group. If all measures were unavailable or could not be calculated from the data and the SOFA score was less than the threshold, the SOFA score was set to “NA” to indicate it could not be calculated. For all remaining patients, an assignment of “low risk of septic shock” was made. The qSOFA score was calculated by accounting for mental status, respiratory rate, and blood pressure as described by Seymour et al.³² If the qSOFA score was 2 or higher, an assignment of “high risk of septic shock” was made. If the qSOFA score was 1 with all 3 measures available in the data set, an assignment of “low risk of septic shock” was made. For the remaining patients, qSOFA-based assignment of septic shock was set to “NA.”

The SIRS score was calculated based on 4 criteria: body temperature, heart rate, respiratory rate, and white blood cell counts.³³ Patients whose scores added up to 2 or greater were considered positive for SIRS (“high risk of septic shock”) while other patients where the score was calculated were negative for

SIRS (“low risk of septic shock”). The Modified Early Warning Score (MEWS)³⁴ was calculated based on the patients systolic blood pressure, heart rate, respiratory rate, temperature, and Alert, Voice, Pain, and Unresponsive (AVPU) scores. The AVPU³⁵ scores which are an approximation of the GCS scores were not directly available in the MIMIC-III database and hence were not estimated. The GCS scores were used in place of AVPU scores based on work by Kyriacos et al.³⁶ The GCS scores were available in the MIMIC-III database. A MEWS score of 2 or greater was used for an assignment of “high risk of septic shock,” whereas a combination of availability of all data and score of less than 2 was used for assignment of “low risk of septic shock.” All other patients were assigned to “NA” category based on MEWS scores.

Risk factor selection for predictive model

All first- and second-degree neighbors of septic shock diagnosis from the cause-and-effect networks were considered for final model selection. For each Bayesian network, stepwise forward selection was performed to identify neighbors that were most predictive of septic shock. Based on manual analysis of the selected factors, lab results that were most predictive of sepsis were identified. Threshold for the general linear model for predicting septic shock was selected to maximize sensitivity and specificity. All variables considered were discrete variables as described earlier in the “Methods” section. A de novo predictive model was built based on lab values that were identified through this method. The possible discrete values for the lab values were “not measured,” “normal,” and “abnormal.” A decision tree was built manually to enhance the interpretation of the regression model.

Results and Discussion

Patient cohorts and novel predictive model for septic shock

A total of 872 septic shock patients and 8293 control patients were identified from the MIMIC-III database. Septic shock patients were identified using the algorithm for digital diagnosis as described in the “Methods” section. The control cohort only contains patients who did not have a digital diagnosis of septic shock during their stay at the ICU. Details of the patient populations are shown in Tables 1 and 2. As shown in Table 2, the patient cohorts have differences in the demographic characteristics. For example, there is a difference in age distribution of patients in the 2 cohorts. It is already known that patient age increases the risk of sepsis.³⁷ In line with this expectation, from Table 2, we can see that the control population tends to be younger. Diagnosis of septic shock is associated with significantly higher mortality rates while in the ICU (relative risk of death in septic shock patients = 2.09) as has been documented in previous studies.^{2,38}

Table 1. Patient cohort for sepsis was identified based on definition as described in Figure 1.

PATIENT COHORT	NO. OF PATIENTS	
	COMPLETE DATA SET	24-HR WINDOW
All	46 520	533
Septic shock	872 (1.87%)	142 (26.6%)
Control	8293	391

A matching but larger control cohort was defined. All analyses were performed on the basis of data per admission.

For 533 patients (391 control cohort patients and 142 septic shock cohort patients), data were available at least 24 to 30 hours prior to diagnosis of septic shock. Other patients did not contain any data earlier than 24 hours prior to diagnosis of septic shock. This cohort of 533 patients (24-hour cohort) was used to infer Bayesian cause-and-effect networks. The data set consisted of interpretations of 48 lab tests, 126 ICD-9 diagnosis code groups, 4 procedure code groups, patient sex, patient age, death during ICU stay, and digital diagnosis status of septic shock. In the 24-hour cohort, a large proportion of the patients that died during the ICU stay had a digital diagnosis of septic shock (43%) comparable with published accounts of causes of ICU mortality.^{39,40}

Figure 2 shows the complete cause-and-effect network and some sample subnetworks. The network after filtering out low-frequency edges consisted of 181 nodes connected by 353 edges. When 2 data features are connected by an edge in the network, it should be interpreted as the upstream feature (demographics, diagnosis, procedure, or lab result) probabilistically leading to the downstream feature (new diagnosis, procedure being performed, or specific lab test status) in a patient. For example, from the inset in Figure 2, it can be inferred that patient age trends with high chances of diagnosis of “coronary atherosclerosis and other heart disease.” This inference is in line with the knowledge that the risk of atherosclerosis and heart disease goes up with age. As a second example from the inset panel, “disorders of lipid metabolism” (which includes codes for high cholesterol levels) is linked to “coronary atherosclerosis and heart disease” as expected. A subnetwork around septic shock is shown in Figure 2 (right panel). Several diagnosis code groups and lab tests are linked to septic shock.

The final regression model that was selected from the selected subnetwork (from Figure 2) is shown in Figure 3. Figure 3 also shows details of the predictive model that was built and its performance characteristics. The model attempts to classify patients in high- or low-risk categories for sepsis based on the status of several lab tests. The lab tests in the predictive model (hemoglobin levels, blood pH, and whole blood potassium levels) were assigned values of not measured and measured (normal or abnormal). A tree representation of the

Table 2. A χ^2 test was performed to compare the patient characteristics between the control and septic shock populations.

PATIENT CHARACTERISTICS	TYPE	COMPLETE COHORT		24-HR COHORT	
		CONTROL	SEPTIC SHOCK	CONTROL	SEPTIC SHOCK
Gender	M	3919 (44.68%)	459 (54.71%)	211 (53.96%)	79 (55.63%)
	F	4853 (55.32%)	380 (45.29%)	180 (46.04%)	63 (44.37%)
Age, y	<18	1200 (13.68%)	28 (3.34%)	0 (0%)	0 (0%)
	18-30	380 (4.33%)	17 (2.03%)	17 (4.35%)	4 (2.82%)
	31-50	1350 (15.39%)	102 (12.16%)	78 (19.95%)	23 (16.2%)
	51-60	1291 (14.72%)	142 (16.92%)	65 (16.62%)	23 (16.2%)
	61-70	1595 (18.18%)	141 (16.81%)	66 (16.88%)	20 (14.08%)
	71-80	1521 (17.34%)	173 (20.62%)	81 (20.72%)	36 (25.35%)
	81-89	1047 (11.94%)	171 (20.38%)	52 (13.3%)	33 (23.24%)
	90+	388 (4.42%)	65 (7.75%)	32 (8.18%)	3 (2.11%)
Length of stay, d	0	95 (1.08%)	116 (13.83%)	0 (0%)	0 (0%)
	1-5	3584 (40.86%)	546 (65.08%)	141 (36.06%)	75 (52.82%)
	6-10	2521 (28.74%)	101 (12.04%)	135 (34.53%)	37 (26.06%)
	11-15	1063 (12.12%)	33 (3.93%)	51 (13.04%)	11 (7.75%)
	16+	33 (0.38%)	2 (0.24%)	64 (16.37%)	19 (13.38%)
Died during ICU stay	Yes	687 (7.83%)	835 (99.52%)	29 (7.42%)	142 (100%)
	No	8085 (92.17%)	4 (0.48%)	362 (92.58%)	0 (0%)

Abbreviation: ICU, intensive care unit.

The null hypothesis that both control and sepsis populations are similar was rejected for all 4 characteristics ($P < .01$). Septic shock patients tend to be older, have a higher proportion of men, and have higher death rates during ICU stay.

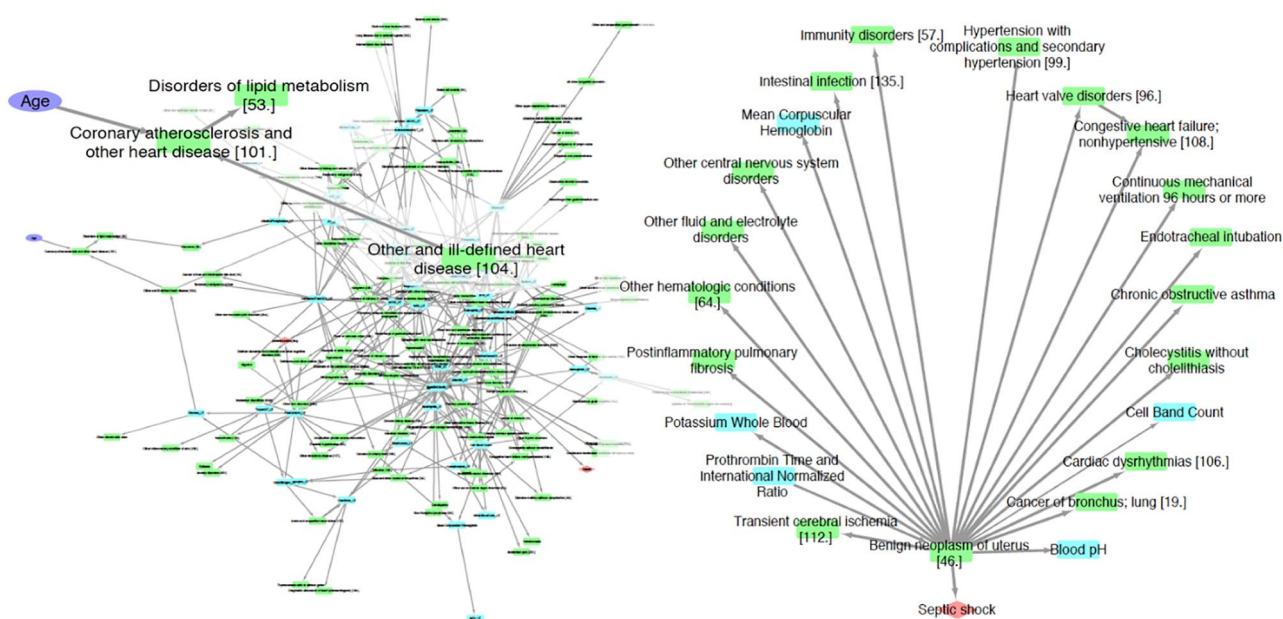


Figure 2. The left panel shows the complete summary network from bAICIS. A zoomed-in view of a portion of the complete network including age has been superimposed on the complete network. The right panel shows the first- and second-degree neighbors of septic shock. Key for nodes in network: purple ellipse—patient demographics, orange diamond—outcomes (death during intensive care unit stay, septic shock), green rectangle—diagnosis and procedure codes, blue rectangle—lab tests.

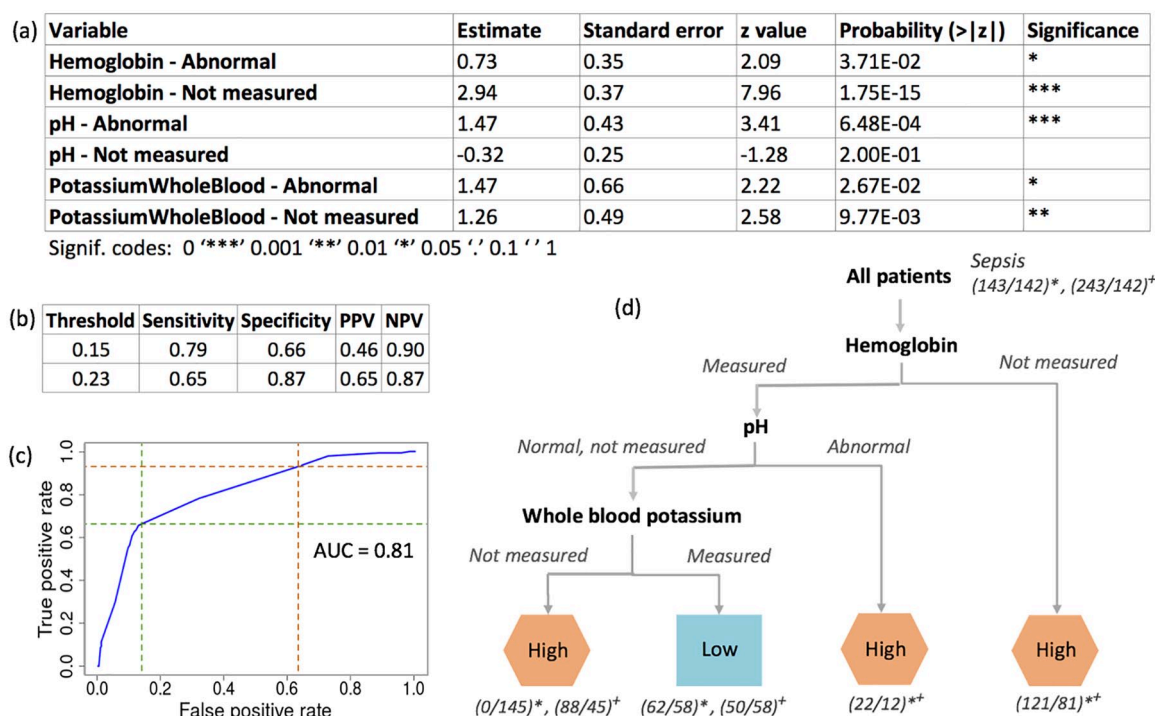


Figure 3. (A) This panel shows the details of the regression model for predicting patients' risk of progressing into septic shock 24 hours prior to diagnosis of septic shock. (B) Two different model cutoffs were selected for assessing model performance. (C) Receiver operating characteristic curve for model. Orange dashed line—0.15 cutoff, green dashed line—0.23 cutoff. The AUC curve was built by interpolating between points measured in the data set as predictive data was discrete. (D) This panel shows a decision tree that represents the regression model used to identify patients at high risk of sepsis. Two different thresholds were selected to allow for different false-negative and true-positive rates. The thresholds are 0.23(*) and 0.15(+). Predictions are made to classify patients as “high” risk and “low” risk. For each prediction, the following information is presented: (number predicted/actual numbers based on assessed septic shock risk). AUC indicates area under the curve; NPV, negative predictive value; PPV, positive predictive value.

regression model is also shown in Figure 3 to enable easy interpretation of the model.

Comparison With current tools for evaluation for septic shock

We attempted to improve current scoring systems such as SOFA by building regression models. The “enhanced” models were defined as follows: “Progression to septic shock”–Scoring method (One of SOFA, qSOFA, SIRS, MEWS) + “lab values from de-novo predictive model.” The “score-only” models only used one of SOFA, qSOFA, SIRS, or MEWS. Performance of the “score-only” models and the “enhanced” models on the MIMIC-III database is shown in Table 3. Only patients for whom both the score calculation was possible and lab data were available for 24-hour time point were included in this analysis. Cohorts for each score were created separately based on data availability. It was observed that while the SOFA score was more complex due to the large number of clinical measures needed, the data were more readily available (279 patients in cohort). Cohorts for other scores were smaller due to larger degree of missingness of relevant data.

Discussion

While several definitions of sepsis (sepsis 1, sepsis 2, and sepsis 3) are in use, they define different patient populations with

differing clinical trajectories. Sepsis 1 and sepsis 2 have been shown in several studies to have poor sensitivity and hence have lower mortality rates^{41–43} when compared with sepsis 3. While, in theory, any definition of sepsis can be used in the development of predictive algorithm, the intended end point of this work was to develop a screening algorithm for septic shock. Sepsis 3 was chosen as a basis of this work due to greater specificity and higher mortality rates in this population depending on patient's responsiveness to fluids, need for vasopressor use, and lactate levels. Using training data with higher specificity is likely to result in an algorithm with better clinical utility as a screening tool for sepsis—identify a small, high-risk patient subpopulation.

It has been shown in several studies that even small delays in diagnosis of sepsis can lead to significant impacts on outcomes^{44–46} in the form of high mortality rates. Therefore, we attempted to build an algorithm to predict patients at risk of progression into septic shock 24 hours ahead of time. Data corresponding to patient status between 24 and 30 hours prior to diagnosis of septic shock were used to create predictive model for septic shock. Many prior attempts have been made to develop such algorithms and early warning has ranged from 4 to 24 hours prior to diagnosis of sepsis.^{11–17} Effective comparison of various algorithms, though, possible will involve major efforts and are outside of the scope of the work presented here.

Table 3. Performance of “score-only” and “enhanced” models.

MODEL	SAMPLE SIZE	MATCH BY	NEGATIVE PREDICTIVE VALUE	POSITIVE PREDICTIVE VALUE	SENSITIVITY	SPECIFICITY
<i>qSOFA</i>	<i>46</i>	—	<i>0.55</i>	<i>0.54</i>	<i>0.55</i>	<i>0.54</i>
Augmented qSOFA model		Sensitivity	0.61	0.64	0.56	0.68
		Specificity	0.66	0.61	0.73	0.54
<i>SOFA</i>	<i>279</i>	—	<i>0.85</i>	<i>0.68</i>	<i>0.12</i>	<i>0.99</i>
Augmented SOFA model		Sensitivity	0.85	0.72	0.12	0.99
		Specificity	0.86	0.69	0.13	0.99
<i>SIRS</i>	<i>57</i>	—	<i>0.99</i>	<i>0.21</i>	<i>0.95</i>	<i>0.51</i>
Augmented SIRS model		Sensitivity	0.99	0.27	0.95	0.64
		Specificity	0.99	0.21	0.95	0.51
<i>MEWS</i>	<i>29</i>	—	<i>0.97</i>	<i>0.21</i>	<i>0.85</i>	<i>0.63</i>
Augmented MEWS model		Sensitivity	0.98	0.24	0.85	0.70
		Specificity	0.98	0.21	0.89	0.63

Abbreviations: MEWS, modified early warning score; qSOFA, quick sequential organ failure assessment; SIRS, systemic inflammatory response syndrome; SOFA, sequential organ failure assessment.

Youden index was used for threshold selection in the “score-only” models. Threshold for the “enhanced” models were calculated by fixing either sensitivity or specificity to that of corresponding “score-only” models. The values in italics were optimal threshold using Youden's index.

Complexity in comparing algorithms arises from multiple causes—variations in definitions of sepsis, reproducibility of methods, differences in evaluation of model performance, and data sets used.

One potential clinical application of a screening tool for septic shock is that it is used to identify a small subset of patients at very high risk. It should be feasible to closely monitor these patients or potentially start early intervention prior to confirmed diagnosis of septic shock. These high-risk patients can be provided antimicrobial treatments and fluids as they are known to be the most effective therapies for septic shock.⁴⁷⁻⁴⁹ Early administration of antibiotics and fluids has been shown to provide most benefit measured through mortality rates.⁵⁰⁻⁵³ Use of antimicrobial treatments and fluid resuscitation is relatively benign, whereas the outcomes can be lethal if treatment is delayed. Given that earlier detection of sepsis is known to result in better outcomes, therefore, use of algorithm can facilitate better allocation of resources to identify patients with sepsis early in the course of the disease. Despite the small number of patients who develop septic shock, this algorithm has moderately high positive predictive value and high negative predictive value as shown in Figure 3. Additional work is necessary to validate the algorithm using a larger or more diverse patient population.

A few different scoring metrics have been created for clinical use to track sepsis in patients in ICU. One of the most well-known scores to assess sepsis is the sepsis-related organ failure assessment (SOFA) score. It was introduced in 1996 by Vincent

et al.⁵⁴ A variant of the SOFA score, qSOFA, is primarily used for mortality prediction but not to clinical track sepsis. Modified early warning score has been used as a tool for screening for sepsis and to identify patients at risk of clinical deterioration.⁵⁵ Systemic inflammatory response syndrome^{33,56} scores are also used to assess patient sepsis status.

From these results shown in Table 3, we observe that while calculation of qSOFA is based only on a few measures and therefore is easy to estimate in the clinic, the model performance in risk stratification of patients is weak as has been observed in several earlier studies.⁵⁷⁻⁵⁹ Only about half the patients at high or low risk of developing septic shock are classified correctly. The qSOFA has been observed to be a better predictor of in-hospital mortality but poorly predictive of severe sepsis.⁶⁰ Based on results of this analysis, we concur with earlier studies that qSOFA score is a poor predictor of severe sepsis or septic shock. The performance of qSOFA could be improved by adding variables from the de novo sepsis prediction algorithm. The SOFA score while having good model performance in correctly classifying risk is based on many more variables and hence is more difficult to calculate in the clinical setting. Based on the implementation of SIRS and MEWS scores, they are observed to have strong performance in identifying patients who are at low risk of septic shock but do not identify a large majority of patients at high risk of septic shock. Some previous work has shown that SIRS and SOFA scores have low sensitivity⁵⁹ validating our observations. The SOFA score could not be significantly improved by augmenting with

the de novo model, whereas the specificity could be improved in the augmented models for SIRS and MEWS.

Conclusions


In this article, we present an algorithm for screening patients at high risk of septic shock and a number of novel risk factors for sepsis that were discovered using data-driven approaches and advanced statistical analysis techniques. The screening tool built as part of this work should be further developed for clinical use where it will be used to screen for patients in need of close monitoring due to their increased risk of developing septic shock. This algorithm was developed on data from one hospital system. Therefore, additional work needs to be done to validate and refine the risk stratification model presented here on a larger multicenter data set.

The novel risk factors are laboratory measurements and are accessible in the hospital setting. The incorporation of the new data into the currently used sepsis prediction models provides a modest improvement of the prediction accuracy. However, even a slight improvement in patient selection could potentially have an enormous financial impact and save patient lives.

Author Contributions

CRY and VV performed the analysis. NRN, VRA, VV conceived the idea of this study and VV wrote the manuscript. NRN, VRA and VV worked in the final approval of the version to be published.

ORCID iD

Vijetha Vemulapalli  <https://orcid.org/0000-0001-6247-2167>

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Fleischmann C, Scherag A, Adhikari NKJ, et al. Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *Am J Respir Crit Care Med*. 2015;193:259-272. doi:10.1164/rccm.201504-0781OC.
2. Martin GS. Sepsis, severe sepsis and septic shock: changes in incidence, pathogenesis and outcomes. *Expert Rev Anti Infect Ther*. 2012;10:701-706. doi:10.1586/eri.12.50.
3. Rhee C, Dantes R, Epstein L, et al. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. *JAMA*. 2017;318:1241-1249. doi:10.1001/jama.2017.13836.
4. Arefian H, Heublein S, Scherag A, et al. Hospital-related cost of sepsis: a systematic review. *J Infect*. 2017;74:107-117. doi:10.1016/j.jinf.2016.11.006.
5. Herasevich V, Afessa B, Chute CG, Gajic O. Designing and testing computer based screening engine for severe sepsis/septic shock. *AMIA Annu Symp Proc*. 2008;2:966.
6. Herasevich V, Pieper MS, Pulido J, Gajic O. Enrollment into a time sensitive clinical study in the critical care setting: results from computerized septic shock sniffer implementation. *J Am Med Inform Assoc*. 2011;18:639-644. doi:10.1136/amiajnl-2011-000228.
7. Semler MW, Weavind L, Hooper MH, et al. An electronic tool for the evaluation and treatment of sepsis in the ICU: a randomized controlled trial. *Crit Care Med*. 2015;43:1595-1602. doi:10.1097/CCM.0000000000001020.
8. Larosa JA, Ahmad N, Feinberg M, Shah M, Dibrienza R, Studer S. The use of an early alert system to improve compliance with sepsis bundles and to assess impact on mortality. *Crit Care Res Pract*. 2012;2012:980369. doi:10.1155/2012/980369.
9. Dellinger RP, Levy MM, Rhodes A, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Med*. 2013;39:165-228. doi:10.1007/s00134-012-2769-8.
10. MacArthur RD, Miller M, Albertson T, et al. Adequacy of early empiric antibiotic treatment and survival in severe sepsis: experience from the MONARCS trial. *Clin Infect Dis*. 2004;38:284-288. doi:10.1086/379825.
11. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Crit Care Med*. 2018;46:547-553. doi:10.1097/CCM.0000000000002936.
12. Liu R, Greenstein JL, Granite SJ, et al. Data-driven discovery of a novel sepsis pre-shock state predicts impending septic shock in the ICU. *Sci Rep*. 2019;9:6145. doi:10.1038/s41598-019-42637-5.
13. Wang S-L, Wu F, Wang B. Prediction of severe sepsis using SVM model. *Adv Exp Med Biol*. 2010;680:75-81. doi:10.1007/978-1-4419-5913-3_9.
14. Desautels T, Calvert J, Hoffman J, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. *JMIR Med Inform*. 2016;4:e28. doi:10.2196/medinform.5909.
15. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. 2018;24:1716-1720. doi:10.1038/s41591-018-0213-5.
16. Lukaszewski RA, Yates AM, Jackson MC, et al. Presymptomatic prediction of sepsis in intensive care unit patients. *Clin Vaccine Immunol*. 2008;15:1089-1094. doi:10.1128/CCI.00486-07.
17. Mao Q, Jay M, Hoffman JL, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open*. 2018;8:e017833. doi:10.1136/bmjopen-2017-017833.
18. Finlay GD, Rothman MJ, Smith RA. Measuring the modified early warning score and the Rothman index: advantages of utilizing the electronic medical record in an early warning system. *J Hosp Med*. 2014;9:116-119. doi:10.1002/jhm.2132.
19. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035. doi:10.1038/sdata.2016.35.
20. Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet—components of a new research resource for complex physiologic signals. *Circulation*. 2000;101:E215-E220.
21. MIMIC-III critical care database. <https://mimic.physionet.org/gettingstarted/access/>. Accessed August 28, 2018.
22. Shabot MM. The HP CareVue clinical information system. *Int J Clin Monit Comput*. 1997;14:177-184. doi:10.1007/BF03356592.
23. Shvachko K, Kuang H, Radia S, Chansler R. The Hadoop distributed file system. In: Khatib MG, He X, Factor M, (eds.) *Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*. MSST '10. Washington, DC: IEEE Computer Society; 2010:1-10. doi:10.1109/MSST.2010.5496972.
24. Thusoo A, Sarma JS, Jain N, et al. Hive: a warehousing solution over a map-reduce framework. *Proc VLDB Endow*. 2009;2:1626-1629.
25. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014.
26. Kadri SS, Rhee C, Strich JR, et al. Estimating ten-year trends in septic shock incidence and mortality in United States academic medical centers using clinical data. *Chest*. 2017;151:278-285. doi:10.1016/j.chest.2016.07.010.
27. Research C for DE and drug approvals and databases—National Drug Code Directory. <https://www.fda.gov/drugs/informationondrugs/ucm142438.htm>. Accessed March 8, 2019.
28. Agency for Healthcare Research Quality: Clinical Classifications Software (CCS) summary download. <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>; <http://www.webcitation.org/78Dsm3Vi>. Accessed March 8, 2019.
29. Vemulapalli V, Qu J, Garren JM, et al. Non-obvious correlations to disease management unraveled by Bayesian artificial intelligence analyses of CMS data. *Artif Intell Med*. 2016;74:1-8. doi:10.1016/j.artmed.2016.11.001.
30. Huang M, Yee CR, Sukhatme VP, Narain NR, Akmaev VR, Vemulapalli V. Identification of novel risk factors for hospital admission after colonoscopy exam by analyzing population wide data with artificial intelligence technology bAICIS®. In: Nagib C, Belkis S, Michael S, Andres T, (eds.) *Proceedings of the 22nd World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2018)*. Winter Garden, FL: International Institute of Informatics and Systemics; Vol.III. 2018:130-133.
31. Jones AE, Trzeciak S, Kline JA. The Sequential Organ Failure Assessment Score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. *Crit Care Med*. 2009;37:1649-1654. doi:10.1097/CCM.0b013e31819def97.
32. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. 2016;315:762-774. doi:10.1001/jama.2016.0288.
33. Bone RC, Balk RA, Cerra FB, et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest*. 1992;101:1644-1655.
34. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified early warning score in medical admissions. *QJM*. 2001;94:521-526. doi:10.1093/qjmed/94.10.521.

35. Kelly CA, Upex A, Bateman DN. Comparison of consciousness level assessment in the poisoned patient using the alert/verbal/painful/unresponsive scale and the Glasgow Coma Scale. *Ann Emerg Med*. 2004;44:108-113. doi:10.1016/j.annemergmed.2004.03.028.
36. Kyriacos U, Jelsma J, James M, Jordan S. Monitoring vital signs: development of a modified early warning scoring (MEWS) system for general wards in a developing country. *PLoS ONE*. 2014;9:e87073. doi:10.1371/journal.pone.0087073.
37. Martin GS, Mannino DM, Moss M. The effect of age on the development and outcome of adult sepsis. *Crit Care Med*. 2006;34:15-21. doi:10.1097/01.ccm.0000194535.82812.ba.
38. Brun-Buisson C, Doyon F, Carlet J. Bacteremia and severe sepsis in adults: a multicenter prospective survey in ICUs and wards of 24 hospitals. *Am J Respir Crit Care Med*. 1996;154:617-624. doi:10.1164/ajrccm.154.3.8810595.
39. Angus DC, Wax RS. Epidemiology of sepsis: an update. *Crit Care Med*. 2001;29:S109-S116.
40. Braber A, van Zanten AR. Unravelling post-ICU mortality: predictors and causes of death. *Eur J Anaesthesiol*. 2010;27:486-490. doi:10.1097/EJA.0b013e3283333aac.
41. Cheng B, Li Z, Wang J, et al. Comparison of the performance between sepsis-1 and sepsis-3 in ICUs in China: a retrospective multicenter study. *Shock*. 2017;48:301-306. doi:10.1097/SHK.0000000000000868.
42. Vincent J-L, Mira J-P, Antonelli M. Sepsis: older and newer concepts. *Lancet Respir Med*. 2016;4:237-240. doi:10.1016/S2213-2600(15)00522-6.
43. Takauji S, Hayakawa M, Fujita S. A nationwide comparison between sepsis-2 and sepsis-3 definition in Japan [published online ahead of print January 13, 2019]. *J Intensive Care Med*. doi:10.1177/0885066618823151.
44. Kumar A, Roberts D, Wood KE, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*. 2006;34:1589-1596. doi:10.1097/01.CCM.00000217961.75225.E9.
45. Ferrer R, Martin-Loeches I, Phillips G, et al. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Crit Care Med*. 2014;42:1749-1755. doi:10.1097/CCM.0000000000000330.
46. Whiles BB, Deis AS, Simpson SQ. Increased time to initial antimicrobial administration is associated with progression to septic shock in severe sepsis patients. *Crit Care Med*. 2017;45:623-629. doi:10.1097/CCM.0000000000000262.
47. Bloos F. Diagnosis and therapy of sepsis. *J Emerg Crit Care Med*. 2018;2. <http://jccm.amegroups.com/article/view/4021>. Accessed July 18, 2019.
48. Polat G, Ugan RA, Cadirci E, Halici Z. Sepsis and septic shock: current treatment strategies and new approaches. *Eurasian J Med*. 2017;49:53-58. doi:10.5152/eurasianjmed.2017.17062.
49. Plevin R, Callcut R. Update in sepsis guidelines: what is really new. *Trauma Surg Acute Care Open*. 2017;2:e000088. doi:10.1136/tsaco-2017-000088.
50. 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;43:304-377. doi:10.1007/s00134-017-4683-6.
51. Ferrer R, Artigas A, Suarez D, et al. Effectiveness of treatments for severe sepsis: a prospective, multicenter, observational study. *Am J Respir Crit Care Med*. 2009;180:861-866. doi:10.1164/rccm.200812-1912OC.
52. Bloos F, Ruddel H, Thomas-Ruddel D, et al. Effect of a multifaceted educational intervention for anti-infectious measures on sepsis mortality: a cluster randomized trial. *Intensive Care Med*. 2017;43:1602-1612. doi:10.1007/s00134-017-4782-4.
53. Howell MD, Davis AM. Management of sepsis and septic shock. *JAMA*. 2017;317:847-848. doi:10.1001/jama.2017.0131.
54. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-Related Organ Failure Assessment) Score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707-710. doi:10.1007/bf01709751.
55. Roney JK, Whitley BE, Maples JC, Futrell LS, Stunkard KA, Long JD. Modified early warning scoring (MEWS): evaluating the evidence for tool inclusion of sepsis screening criteria and impact on mortality and failure to rescue. *J Clin Nurs*. 2015;24:3343-3354. doi:10.1111/jocn.12952.
56. Levy MM, Fink MP, Marshall JC, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference. *Crit Care Med*. 2003;31:1250-1256. doi:10.1097/01.CCM.0000050454.01978.3B.
57. Dorsett M, Kroll M, Smith CS, Asaro P, Liang SY, Moy HP. qSOFA has poor sensitivity for prehospital identification of severe sepsis and septic shock. *Prehosp Emerg Care*. 2017;21:489-497. doi:10.1080/10903127.2016.1274348.
58. Askim Å, Moser F, Gustad LT, et al. Poor performance of quick—SOFA (qSOFA) score in predicting severe sepsis and mortality—a prospective study of patients admitted with infection to the emergency department. *Scand J Trauma Resusc Emerg Med*. 2017;25:56. doi:10.1186/s13049-017-0399-4.
59. Tugul S, Carron P-N, Yersin B, Calandra T, Dami F. Low sensitivity of qSOFA, SIRS criteria and sepsis definition to identify infected patients at risk of complication in the prehospital setting and at the emergency department triage. *Scand J Trauma Resusc Emerg Med*. 2017;25:108. doi:10.1186/s13049-017-0449-y.
60. Raith EP, Udy AA, Bailey M, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA*. 2017;317:290-300. doi:10.1001/jama.2016.20328.