

---

## Research and Applications

# Sepsis surveillance: an examination of parameter sensitivity and alert reliability

Robert C. Amland, Mark Burghart, and J. Marc Overhage

Population Health, Cerner Corporation, Kansas City, Missouri, USA

Corresponding Author: Robert C. Amland, PhD, Cerner Corporation, 2800 Rockcreek Parkway, Kansas City, MO 64117, USA (bob.amland@cerner.com); 8854 Pine Street, Lenexa, KS 66220, USA (bobamland@gmail.com)

Received 13 December 2018; Revised 18 March 2019; Editorial Decision 23 April 2019; Accepted 26 April 2019

### ABSTRACT

**Objective:** To examine performance of a sepsis surveillance system in a simulated environment where modifications to parameters and settings for identification of at-risk patients can be explored in-depth.

**Materials and Methods:** This was a multiple center observational cohort study. The study population comprised 14 917 adults hospitalized in 2016. An expert-driven rules algorithm was applied against 15.1 million data points to simulate a system with binary notification of sepsis events. Three system scenarios were examined: a scenario as derived from the second version of the Consensus Definitions for Sepsis and Septic Shock (SEP-2), the same scenario but without systolic blood pressure (SBP) decrease criteria (near SEP-2), and a conservative scenario with limited parameters. Patients identified by scenarios as being at-risk for sepsis were assessed for suspected infection. Multivariate binary logistic regression models estimated mortality risk among patients with suspected infection.

**Results:** First, the SEP-2-based scenario had a hyperactive, unreliable parameter SBP decrease >40 mmHg from baseline. Second, the near SEP-2 scenario demonstrated adequate reliability and sensitivity. Third, the conservative scenario had modestly higher reliability, but sensitivity degraded quickly. Parameters differed in predicting mortality risk and represented a substitution effect between scenarios.

**Discussion:** Configuration of parameters and alert criteria have implications for patient identification and predicted outcomes.

**Conclusion:** Performance of scenarios was associated with scenario design. A single hyperactive, unreliable parameter may negatively influence adoption of the system. A trade-off between modest improvements in alert reliability corresponded to a steep decline in condition sensitivity in scenarios explored.

**Key words:** sepsis, classification, expert systems, decision support, ergonomics

---

## BACKGROUND AND SIGNIFICANCE

Despite increased awareness of sepsis, accurately identifying patients at risk remains a challenge.<sup>1</sup> This phenomenon is applicable to sepsis surveillance systems, noteworthy for having characteristic heterogeneous designs. However, commonalities across systems typically include a limited number of parameters and alert criteria threshold settings selected to attain high reliability at a detrimental cost of declining sensitivity.<sup>2–4</sup> A countervailing approach is a broadening of parameter inclusion applicable to sepsis surveillance, which could

play a larger role in promoting early recognition and treatment, and improving the likelihood of survival.<sup>5</sup>

A sepsis surveillance system is typically available in real-time at the point of care and integrates clinical decision support functionality into the clinical workflow.<sup>6</sup> The system continuously screens a patient's physiological data and is capable of delivering a notification to a provider within minutes of activation.<sup>7–9</sup> Screening typically begins when the patient presents to the emergency department or is otherwise admitted to the hospital, and runs uninterrupted until discharge.<sup>10,11</sup>

**KEY FINDINGS**

- Omission error can be minimized by establishing an adequate reliability metric for adoption, with an understanding that performance of parameters is not homogenous.
- Prior to modifying system parameter configurations, realize that small improvements in reliability may be offset by a corresponding steep decline in condition sensitivity.
- Constraining systems by design promotes substitution effects on parameters that increase mortality risk for a sizable subgroup of patients.

Although many surveillance systems have good performance characteristics, none are 100% accurate. The system's usefulness is premised upon a provider's subjective perception of reliability, which in itself is associated with the design of system parameters and alert criteria thresholds.<sup>12</sup> A reliable alert reduces harmful cry-wolf effects, such as ignoring alert notifications or delays in response.<sup>13</sup> Moreover, a surveillance system characterized by an adequate alert reliability (eg,  $\geq 70\%$  positive predictive value [PPV]) increases the likelihood of response to alert notifications and attaining protocol compliance.<sup>14</sup> In contrast, providers' reliance on the system to accurately screen-in at-risk patients is supported by systems possessing high condition sensitivity. Nevertheless, an intentionally sensitive high-reliance design may render the system prone to false positives and an unacceptably low reliability metric, a trade-off which has implications for adoption and use.<sup>15-17</sup>

Two types of error tensions are inherent in surveillance systems. First, commission error (ie, over-diagnosing or treatment despite the absence of the condition; false positive) is observable when an alert activates and delivers a notification to a provider who responds despite the alert being in error. Second, an omission error (ie, missed detection and nonresponse; false negative) occurs when a patient is condition positive, but the system remains silent and a provider does not place an order. Designers of surveillance systems, therefore, may rely upon evidence and heuristics to increase reliability to reduce commission errors while hedging condition sensitivity given an uncertain omission error potential.<sup>18</sup>

**OBJECTIVES**

The objective of this study was to examine performance of a sepsis surveillance system in a simulated environment, where modifications to parameters and alert criteria thresholds to accurately screen-in at-risk patients could be explored. Although early guidance suggested surveillance system performance can be improved substantially by manipulating parameters and alert criteria thresholds, a small adjustment may have an unintentional, detrimental impact on condition sensitivity and become an exercise in frustration.<sup>19</sup> Moreover, the impact of excluding parameters on performance when compared to a system as derived from the second version of the Consensus Definitions for Sepsis and Septic Shock (SEP-2) is generally unknown.<sup>20</sup>

**MATERIALS AND METHODS****Study design**

This was a retrospective multiple center observational cohort study. The study site was eight hospitals located in two different regions in southwest USA. All facilities had an enterprise electronic health record (EHR) system (Millennium: Cerner Corporation, Kansas City,

MO, USA). This study was approved with waiver of informed consent by the Western Institutional Review Board.

The study population included 14 917 patients ( $\geq 18$  years) admitted to the hospital and discharged during a 90-day observation period in 2016. Data were retrieved from an enterprise data warehouse and production system. To account for updates to physiological data, a counter-like key written in SQL was instituted during cleaning and processing of data. The counter-like key was used to sort multiple instances of updated physiological data related to a clinical event. Essentially, this data engineering process allowed for a proper temporal sequencing of many instances related to a single clinical event and important to classification of physiological data. To merge data extracts, a common unique encounter identifier created by the EHR system was available as the primary key.

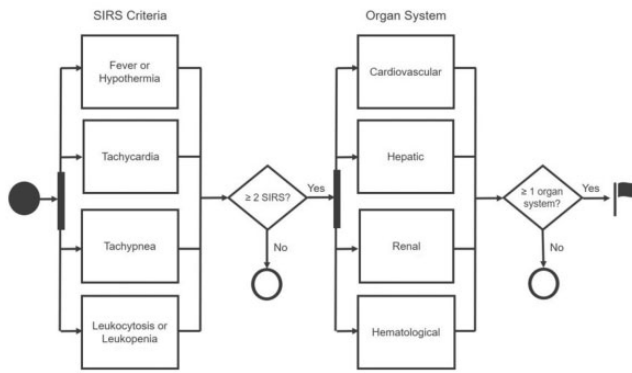
All patient encounters were examined for suspicion of infection, which we defined as microbiology cultures as having been drawn and intravenous (IV) antibiotics administered.<sup>21,22</sup> To identify patients at-risk for sepsis, an expert-driven rules algorithm was developed to simulate a sepsis surveillance system with a binary alert notification. Figure 1 illustrates the foundation model with three scenarios for classification. The SEP-2-based scenario of the model defined the universe of physiological data available for classification, which comprised 15.1 million data points.

The SEP-2-based scenario adhered closely to the definition of severe sepsis recognized by the Centers for Medicare & Medicaid Services.<sup>23</sup> Each scenario thereafter introduced modifications to parameters and/or adjustments to alert criteria thresholds to potentially improve the trade-off between alert reliability and condition sensitivity. The near SEP-2 scenario excluded the parameter corresponding to a decline in systolic blood pressure (SBP)  $>40$  mmHg from the SEP-2-based scenario because some patients may have experienced adrenergic stress-induced elevation in blood pressure, and if normalized, could erroneously activate a severe sepsis alert.<sup>24</sup> In addition to excluding the parameter SBP decrease  $>40$  mmHg, the conservative scenario modified four parameters: an increase of 0.5 mg/dL over 72 h for creatinine; excluded bilirubin  $>10$  mg/dL; and eliminated platelets and international normalized ratio (INR) because these parameters may also indicate primary liver injury.<sup>25-27</sup>

**Definitions**

The primary outcome was identification of patients with severe sepsis, defined as suspected or confirmed infection with clinical evidence of two Systemic Inflammatory Response Syndrome (SIRS) criteria and at-least one organ system dysfunction. A secondary outcome, mortality, was defined as in-hospital death or referral to hospice at discharge.

As illustrated in Figure 1, SIRS was indicated when two of the following four criteria were satisfied: (1) temperature  $>38.3^{\circ}\text{C}$  or  $<36^{\circ}\text{C}$ ; (2) heart rate  $>90$  beats/min; (3) respiratory rate  $>20$  breaths/min; and (4) white blood cell count  $>12\ 000$  cells/ $\text{mm}^3$ ,



Parameter	SEP-2 based scenario	Near SEP-2 scenario	Conservative Scenario
Systolic Blood Pressure < 90 mm Hg or MAP < 65 mm Hg or SBP $\Delta$ ↓ > 40 mm Hg from baseline	●	●	●
Lactate > 2.0 mmol/L	●	●	●
Bilirubin $\geq$ 2.0 mg/dL	●	●	● <sup>1</sup>
Creatinine $\geq$ 2.0 mg/dL	●	●	● <sup>2</sup>
Platelet count < 100,000 $\mu$ L <sup>-1</sup>	●	●	○
International Normalized Ratio > 1.5	●	●	○

Notes. Parameters: ● included or ○ excluded. Threshold: (1) bilirubin  $\geq$  2.0 mg/dL AND  $\leq$  10.0 mg/dL; (2) creatinine  $\Delta$ ↑ 0.5 mg/dL over 72 hours.

**Figure 1.** Sepsis surveillance model and corresponding organ system parameters by surveillance system scenario. SBP: systolic blood pressure; SEP-2: Consensus Definitions for Sepsis and Septic Shock; SIRS: Systemic Inflammatory Response Syndrome; MAP: mean arterial pressure.

<4000 cells/mm<sup>3</sup>, or >10% immature (band) forms. Severe sepsis screen-in was established when  $\geq 2$  SIRS criteria were present, and  $\geq 1$  of the following four organ system dysfunction criteria were satisfied: (1) SBP <90 mm Hg, or mean arterial pressure <65 mm Hg, or SBP decrease >40 mm Hg from baseline, or serum lactate >2.0 mmol/L; (2) total bilirubin:  $\geq 2.0$  mg/dL; (3) serum creatinine:  $\geq 2.0$  mg/dL; and (4) platelet count <100 000  $\mu$ L<sup>-1</sup> or INR >1.5. Organ dysfunction was indicated by itemized parameter criteria by system scenario shown in Figure 1.

Classification of parameters for each system scenario applied the following lookback periods: 12 h for lactate, 72 h for creatinine, and 30 h for the remaining criteria. The first SBP reading within the 30-h lookback period was established as the baseline SBP for the parameter SBP decrease >40 mm Hg from baseline.

Suspected infection was defined by the combination of microbiology cultures drawn and anti-infective antibiotics administered. Microbiology cultures included blood, body fluid, bronchial, catheter tip, cerebrospinal fluid, fungal, ova and parasites, sputum, stool, tissue, urine, and wound. Antibiotics included ampicillin-sulbactam, azithromycin, cefepime, ceftriaxone, ciprofloxacin, clindamycin, fluconazole, fluticasone-salmeterol, levofloxacin, meropenem, piperacillin-tazobactam, and vancomycin.

Demographics and severity of illness measures were calculated at the time of patient arrival to the hospital. The National Early Warning (NEWS) composite acuity score was calculated and categorized by using the first vital signs and neurologic assessment documented, to include respiratory rate, oxygen saturation, temperature, SBP, heart rate, and level of consciousness.<sup>28,29</sup> An indication of severe electrolyte abnormality and metabolic disturbance was calculated using the *apparent* Strong Ion Difference (SIDa), where corrected SIDa = [(Na<sup>+</sup> + K<sup>+</sup> + 1.85) - Cl<sup>-</sup>]; and flagged when SIDa  $\leq 34.0$  or  $\geq 48.0$  mmol/L.<sup>30-33</sup>

**Statistical analysis**

Data were retrospectively analyzed. A confusion matrix, noteworthy for its specific 2 by 2 table layout, was created to calculate performance metrics (ie, PPV and sensitivity) for each system scenario on suspected infection. Multivariate binary logistic regression (MLR) modeling was used to estimate mortality risk among patients with suspected infection. Adjusted odds ratios with 95% confidence intervals were reported for each parameter by system scenario and illustrated in a forest plot. For the parameter included in each MLR model, the absence of that parameter was the reference value. MLR model covariates included patient age, sex, and NEWS acuity cate-

gory. Classification of parameters was coded in Python (version 3.6) using Scikit-learn (version 0.19.2) and Pandas (version 21.1) libraries. The analysis of statistical data was conducted in SPSS v25 (IBM, Inc., Armonk, NY, USA).

**RESULTS**

Classification output merged into the study population analytic set identified 3862 patients who met criteria associated with the SEP-2-based scenario. Characteristics of these patients are shown in Table 1. Less than two-thirds ( $n = 2472$  of 3862; 64%) of patients had a suspicion of infection, with lactate measured among 81% ( $n = 1999$  of 2472) of patients. In comparison, a relative 19% [(3142-3862)/3862] fewer patients and relative 33% [(2598-3862)/3862] fewer patients were identified by the near SEP-2 scenario or the conservative scenario, respectively. Performance of system scenarios showed alert reliability of 64%, 68%, and 71% PPV on suspected infection, while sensitivity was 100%, 87%, and 75% for the SEP-2-based scenario, the near SEP-2 scenario, and the conservative scenario, respectively.

The typical patient with suspected infection was 66 years old and slightly less likely to be female. Upon arrival to the hospital, approximately 5% patients had severe electrolyte abnormality and metabolic disturbance; and one in four patients had NEWS  $\geq 7$  points of which half of them had NEWS  $\geq 9$  points, which suggested a progressive physiologic deterioration. Between 17% and 21% patients either expired in-hospital or referred to hospice at discharge.

Table 2 describes the occurrence of patients by classification and proportion with suspected infection, which ranged from 50% to 85%. The parameter SBP decrease >40 mm Hg from baseline was responsible for identifying the greatest number of patients in the SEP-2-based scenario and demonstrated low reliability on suspected infection and therefore drove a large proportion of the false positive alerts. Classification of parameters on the other two scenarios showed not only a notable relative decline in detected patients as the scenarios became increasingly constrained, but also a complex substitution effect on parameters between scenarios.

In the SEP-2-based scenario, 66% ( $n = 2561$  of 3862) patients had a cardiovascular system parameter indicated compared to 47% ( $n = 1483$  of 3142) patients in the near SEP-2 scenario and 77% ( $n = 2046$  of 2598) patients in the conservative scenario. Moreover, 5% ( $n = 209$ ) patients had a hepatic or renal system parameter indicated in the SEP-2-based scenario compared to 15% ( $n = 477$ ) and

**Table 1.** Characteristics of patients with severe sepsis flag by system scenario

Characteristics	SEP-2-based scenario N (%)	Near SEP-2 scenario N (%)	Conservative scenario N (%)
Hospitalizations	3862 (100)	3142 (100)	2598 (100)
Suspected infection	2472 (64)	2147 (68)	1843 (71)
Lactates measured	1999 (81)	1797 (84)	1591 (86)
Demographics			
Age (y), median (IQR)	66 (52–76)	66 (53–77)	65 (52–76)
Female gender	1173 (48)	995 (46)	854 (46)
First clinical results			
SIDa $\leq 34$ or $\geq 48$ mmol/L	128 (05)	125 (06)	105 (06)
NEWS composite score			
0–4 points	1469 (59)	1243 (58)	1016 (55)
5–6	465 (19)	418 (19)	371 (20)
7–8	301 (12)	260 (12)	242 (13)
9–25	237 (10)	226 (11)	214 (12)
Clinical outcomes			
Expired or hospice	424 (17)	404 (19)	380 (21)
ICU, expired, hospice	848 (34)	787 (37)	734 (40)
LOS (d), median (IQR)	4 (3–9)	5 (3–9)	5 (3–9)

Note: Suspected infection elements include microbiology cultures drawn and anti-infective intravenous antibiotics given. ICU admission within 48 h after arrival.

Abbreviations: ICU: intensive care unit; IQR: interquartile range; LOS: length of stay; NEWS: National Early Warning Score; SEP-2: Consensus Definitions for Sepsis and Septic Shock; SIDa: *apparent* Strong Ion Difference.

**Table 2.** Patients with severe sepsis flag and suspicion of infection

Parameter	SEP-2-based scenario		Near SEP-2 scenario		Conservative scenario	
	N	n (%)	N	n (%)	N	n (%)
SBP $\Delta$ ↓	1083	577 (53)	–	–	–	–
SBP or MAP	719	435 (61)	771	447 (58)	1185	745 (63)
Lactate	759	643 (85)	712	596 (84)	861	729 (85)
Bilirubin	121	69 (57)	120	66 (55)	222	136 (61)
Creatinine	88	44 (50)	357	258 (72)	174	105 (60)
MODS	394	283 (72)	522	383 (73)	156	128 (82)
Platelets	278	185 (67)	255	168 (66)	–	–
INR	420	236 (56)	405	229 (57)	–	–

Note: Parameter SBP  $\Delta$ ↓: SBP decrease  $>40$  mm Hg from baseline.

Abbreviations: INR: international normalized ratio; MAP: mean arterial pressure; MODS: multiple organ dysfunction; SBP: systolic blood pressure; SEP-2: Consensus Definitions for Sepsis and Septic Shock.

15% ( $n = 396$ ) patients in the near SEP-2 scenario and conservative scenario, respectively. Finally, 10% ( $n = 394$ ) patients had multiple organ dysfunction (MODS) indicated in the SEP-2-based scenario compared to 17% ( $n = 522$ ) patients in the near SEP-2 scenario and 6% ( $n = 156$ ) patients in the conservative scenario.

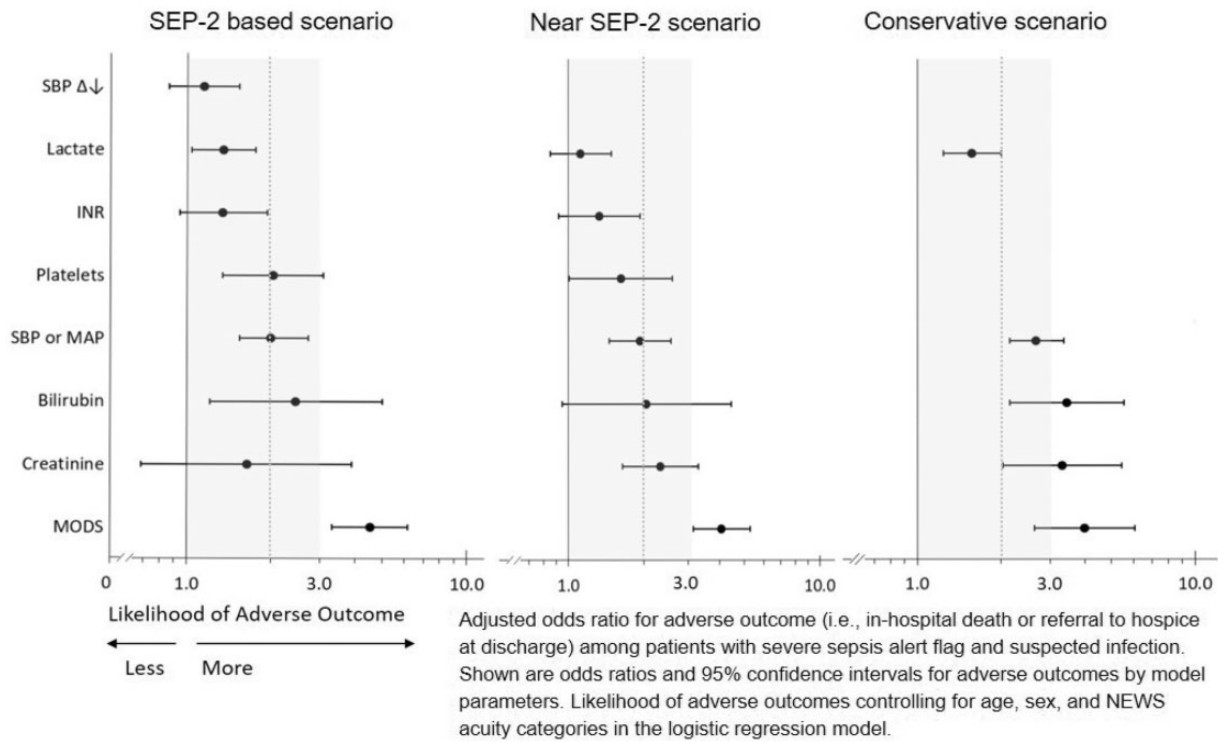
### Substitution effects on parameters among patients with suspected infection

**Membership in both SEP-2-based scenario and near SEP-2 scenario**  
Approximately 87% patients identified by the SEP-2-based scenario who had a suspicion of infection were also identified by the near SEP-2 scenario. Among these patients, one in four had a different classification in the near SEP-2 scenario than their initial classification in the SEP-2-based scenario ( $n = 540$  of 2147; 25%). Of the 540-patient subgroup with a change in classification, the new parameter included 27% cardiovascular, 1% hepatic, 40% renal, 5% hematological, and 27% MODS. The crude mortality rate was 22% ( $n = 118$  of 540) versus 18% ( $n = 286$  of 1607) in the 540-patient

subgroup compared to the other patient subgroup, respectively. When compared with patients who had the same classification between system scenarios, patients with a new classification in the near SEP-2 scenario had 29% increased mortality risk (unadjusted odds ratio 1.29, 95% confidence interval 1.02–1.64).

### Membership in both SEP-2-based scenario and conservative system

Approximately 75% patients identified by the SEP-2-based scenario who had a suspicion of infection were also identified by the conservative scenario. Among these dual membership patients, one in three had a different classification in the conservative scenario than their initial classification in the SEP-2-based scenario ( $n = 646$  of 1843; 35%). Of the 646-patient subgroup with a change in classification, the new parameter included 74% cardiovascular, 11% hepatic, 10% renal, and 5% MODS. The crude mortality rate was 26% ( $n = 165$  of 646) versus 18% ( $n = 215$  of 1197) in the 646-patient subgroup compared to the other patient subgroup. When compared with patients who had the same classification between scenarios,



**Figure 2.** Forest plot of organ system parameters on mortality by surveillance system scenario. INR: international normalized ratio; MAP: mean arterial pressure; MODS: multiple organ dysfunction; NEWS: National Early Warning Score; SBP: systolic blood pressure; SEP-2: Consensus Definitions for Sepsis and Septic Shock.

patients with a new classification in the conservative scenario had 57% increased mortality risk (unadjusted odds ratio 1.57, 95% confidence interval 1.24–1.97).

**Mortality outcomes**

In multivariable analysis, risk-adjusted mortality outcomes among patients with suspected infection differed by classification of parameter. Forest plots illustrated in Figure 2 illuminate mortality risk by parameter within respective system scenarios. An increase in mortality risk was pronounced by the parameter creatinine when comparing the near SEP-2 scenario to the SEP-2-based scenario. Moreover, mortality risk increased for each parameter, with exception of the parameter MODS, when comparing the conservative scenario to the SEP-2-based scenario.

**DISCUSSION**

Finding an agreeable trade-off between reliability and sensitivity in a sepsis surveillance system can be challenging, especially when the objective is to achieve a target performance metric of 70% PPV on suspected infection. The simulation illustrated the SEP-2-based scenario experienced performance issues pertinent to a hyperactive, unreliable parameter SBP decrease >40 mmHg from baseline. By excluding this parameter, a corresponding substitution effect increased alert reliability modestly but lowered sensitivity more dramatically. Mortality risk increased significantly as system configurations became more conservative. This said, the near SEP-2 scenario achieved an adequate trade-off between alert reliability and condition sensitivity particularly when considering mortality risk.

This study is potentially the first study to simulate a sepsis surveillance system that applied scenarios to gain insights into alert ac-

tivation, performance, and mortality risk. In this regard, the study’s findings offer a unique contribution to the literature. The study, nevertheless, has limitations to consider. This was a multiple center observational cohort study involving eight hospitals in the USA. A classification of parameters using structured clinical data occurred 6–9 months after launch of the hospitals’ sepsis management programs, which may have introduced informed presence or other selection bias associated with real-world clinical practice.<sup>34</sup> The accuracy and timeliness of clinical documentation at the patient bedside were not examined; however, multiple instances of clinical results for a same order were identified retrospectively while cleaning and processing of physiological data. Therefore, a counter-like key was coded and instituted when processing data to establish the proper temporal sequence of updated clinical results. Simulation results may not be applicable to other hospitals or health systems.

Sepsis surveillance is an effective approach for reducing mortality risk among at-risk patients.<sup>14,35,36</sup> This study’s findings support the enduring robustness of severe sepsis definition articulated in policy,<sup>23</sup> although room for improving precision exists. Despite being on solid ground when referencing a well-established sepsis surveillance model, a system’s clinical effectiveness may be questioned when a hyperactive unreliable parameter exists on the one-hand, or a more conservatively parameterized design prevails. One aberrant parameter may be enough to produce a negative perception of the whole system. In contrast, a conservatively parameterized system can increase reliability and garner positive perceptions of effectiveness in the near-term, but good-willed perceptions over-time may result in incremental changes that degrade performance.<sup>9</sup> In this regard, condition sensitivity can be improved at a cost of reducing alert reliability in a conservative system, with full-realization that options to improve performance are already limited before creating negative perceptions.



In this light, a SEP-2-based scenario was the intended target system. But, study findings showed a hyperactive parameter SBP decrease >40 mmHg from baseline and had inadequate reliability on suspected infection. The issue pertained to the difficulty in defining a rule for accurate classification of hypotension, considering the existence of substantial variability in time intervals for SBP readings coupled with a myriad of possibilities for an SBP decrease >40 mmHg. This said, the system's reliability improved to an adequate level for response after excluding that parameter. Further examination on the intersection of management of high blood pressure and SIRS is necessary before reintroducing the parameter SBP decrease >40 mmHg from baseline into a sepsis surveillance system.

Several machine learned models to identify patients at-risk for sepsis in the intensive care unit have been published.<sup>37-39</sup> Models have been developed using a variety of definitions of severe sepsis,<sup>20,40</sup> and demonstrated a similar sensitivity (ie, 80–85%) to the near SEP-2 scenario but a lower alert reliability (ie, 20–30% PPV). The near SEP-2 scenario coupled with a clinical suspicion of infection may be a reliable outcome measure to use for machine learning training and testing.<sup>41</sup>

Future research might include model development using the near SEP-2 scenario with suspected infection as a primary target as well as secondary targets defined by the seven organ dysfunction parameters. Given surveillance systems are not 100% accurate, investigators should examine system performance driven by a machine learning-based algorithm coupled with a backstop expert-driven rules algorithm.

## CONCLUSION

Performance of the sepsis surveillance system was associated with its design. A substitution effect on parameters occurred among a subgroup of patients as systems became more conservative. Mortality risk increased significantly when systems applied constraints. This phenomenon suggests the likelihood of survival improves when a sepsis surveillance system has sufficiently adequate reliability that minimizes omission error.

## DATA AVAILABILITY

Data available from the Dryad Digital Repository: <https://dx.doi.org/10.5061/dryad.vh380vb>.

## CONTRIBUTORS

All authors contributed to the study design, analysis, or interpretation of data; R.C.A. and M.B. were responsible for data acquisition. All authors participated in drafting the manuscript for important intellectual content. All authors approved the final version of the manuscript.

## ACKNOWLEDGEMENTS

We appreciate the help given by Erin Hoemann and Leo Perez on data engineering, and Bharat B. Sutariya, MD, on clinical intelligence.

*Conflict of interest statement.* The authors are employed by Cerner Corporation, developer of the Millennium electronic health record system.

## REFERENCES

1. Moskowitz A, Andersen LW, Cocchi M, *et al.* The misapplication of severity of illness scores toward clinical decision making. *Am J Respir Crit Care Med* 2016; 194 (3): 256–8.
2. Alberto L, Marshall AP, Walker R, *et al.* Screening for sepsis in general hospitalized patients: a systematic review. *J Hosp Infect* 2017; 96 (4): 305–15.
3. Despina LA. Automated detection of sepsis using electronic medical record data: a systematic review. *J Healthc Qual* 2017; 39 (6): 322–33.
4. Fernando SM, Tran A, Taljaard M, *et al.* Prognostic accuracy of the Quick Sequential Organ Failure Assessment for mortality in patients with suspected infection: a systematic review and meta-analysis. *Ann Intern Med* 2018; 168 (4): 266–75.
5. Capan M, Hoover S, Ivy JS, *et al.* Not all organ dysfunctions are created equal: prevalence and mortality in sepsis. *J Crit Care* 2018; 48: 257–62.
6. Osheroff JA, Teich JM, Middleton B, *et al.* A roadmap for national action on clinical decision support. *J Am Med Inform Assoc* 2007; 14 (2): 141–5.
7. Umsheid CA, Betesh J, VanZandbergen C, *et al.* Development, implementation, and impact of an automated early warning and response system for sepsis. *J Hosp Med* 2015; 10 (1): 26–31.
8. Amland RC, Lyons JJ, Greene TL, *et al.* A two-stage clinical decision support system for early recognition and stratification of patients with sepsis: an observational cohort study. *J R Soc Med Open* 2015; 6 (10): 1–10.
9. Austrian JS, Jamin CT, Doty GR, *et al.* Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. *J Am Med Inform Assoc* 2018; 25 (5): 523–9.
10. Amland RC, Hahn-Cover KC. Clinical decision support for early recognition of sepsis. *Am J Med Qual* 2016; 31 (2): 103–10.
11. Khurana HS, Groves RH, Simons MP, *et al.* Real-time automated sampling of electronic medical records predicts hospital mortality. *Am J Med* 2016; 129 (7): 688–98.
12. Meyer J, Wiczorek R, Gunzler T. Measures of reliance and compliance in aided visual scanning. *Hum Factors* 2014; 56 (5): 840–9.
13. Parasuraman R, Wickens CD. Humans: still vital after all these years of automation. *Hum Factors* 2008; 50 (3): 511–20.
14. Amland RC, Sutariya BB. An investigation of sepsis surveillance and emergency treatment on patient mortality outcomes: an observational cohort study. *JAMIA Open* 2018; 1 (1): 107–14.
15. Meyer J. Conceptual issues in the study of dynamic hazard warnings. *Hum Factors* 2004; 46 (2): 196–204.
16. Wiczorek R, Manzey D. Supporting attention allocation in multitask environments: effects of likelihood alarm systems on trust, behavior, and performance. *Hum Factors* 2014; 56 (7): 1209–21.
17. Brandt BN, Gartner AB, Moncure M, *et al.* Identifying severe sepsis via electronic surveillance. *Am J Med Qual* 2015; 30 (6): 559–65.
18. Dixon SR, Wickens CD, McCarley JS. On the independence of compliance and reliance: are automation false alarms worse than misses? *Hum Factors* 2007; 49 (4): 564–72.
19. Meyer J, Sheridan TB. The intricacies of user adjustments of alerting thresholds. *Hum Factors* 2017; 59 (6): 901–10.
20. Levy MM, Fink MP, Marshall JC, *et al.* 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Crit Care Med* 2003; 31 (4): 1250–6.
21. 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 (8): 762–74.
22. Levy MM, Gesten FC, Phillips GS, *et al.* Mortality changes associated with mandated public reporting for sepsis: the results of the New York State Initiative. *Am J Respir Crit Care Med* 2018; 198 (11): 1406–12.
23. Centers for Medicare and Medicaid Services, US Department of Health and Human Services. Proposal for Sepsis Voluntary Reporting Measures, SEP-1. [https://www.qualityforum.org/Projects/i-m/Infectious\\_Disease\\_Endorsement\\_Maintenance\\_2012/0500.aspx](https://www.qualityforum.org/Projects/i-m/Infectious_Disease_Endorsement_Maintenance_2012/0500.aspx) Accessed February 11, 2019; subsection 2a1.7 Denominator Details.
24. Dunser MW, Hasibeder WR. Sympathetic overstimulation during critical illness: adverse effects of adrenergic stress. *J Intensive Care Med* 2009; 24 (5): 293–316.

25. Linder A, Fjell C, Levin A, *et al.* Small acute increases in serum creatinine are associated with decreased long-term survival in the critically ill. *Am J Respir Crit Care Med* 2014; 189 (9): 1075–81.
26. Chand N, Sanyal AJ. Sepsis-induced cholestasis. *Hepatology* 2007; 45 (1): 230–41.
27. Stravitz RT, Ellerbe C, Durkalski V, *et al.* Thrombocytopenia is associated with multi-organ system failure in patients with acute liver failure. *Clin Gastroenterol Hepatol* 2016; 14 (4): 613–20.
28. Smith GB, Prytherch DR, Meredith P, *et al.* The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013; 84 (4): 465–70.
29. Corfield AR, Lees F, Zealley I, *et al.* Utility of a single early warning score in patients with sepsis in the emergency department. *Emerg Med J* 2014; 31 (6): 482–7.
30. Kellum JA, Kramer DJ, Pinsky M. Strong Ion Gap: a methodology for exploring unexplained anions. *J Crit Care* 1995; 10 (2): 51–5.
31. Noritomi DT, Soriano FG, Kellum JA, *et al.* Metabolic acidosis in patients with severe sepsis and septic shock: a longitudinal quantitative study. *Crit Care Med* 2009; 37 (10): 2733–9.
32. Egi M, Kim I, Nichol A, *et al.* Ionized calcium concentration and outcome in critical illness. *Crit Care Med* 2011; 39 (2): 314–21.
33. Moskowitz A, Lee J, Donnino MW, *et al.* The association between admission magnesium concentrations and lactic acidosis in critical illness. *J Intensive Care Med* 2016; 31 (3): 187–92.
34. Goldstein BA, Navar AM, Pencina MJ, *et al.* Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *J Am Med Inform Assoc* 2017; 24 (1): 198–208.
35. Amland RC, Haley JM, Lyons JJ. A multidisciplinary sepsis program enabled by a two-stage clinical decision support system: factors that influence patient outcomes. *Am J Med Qual* 2016; 31 (6): 501–8.
36. Westra B, Landman S, Yadav P. Secondary analysis of an electronic surveillance system combined with multi-focal interventions for early detection of sepsis. *Appl Clin Inform* 2017; 8 (1): 46–66.
37. 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 (3): e28.
38. Henry KE, Hager DN, Pronovost PJ, *et al.* A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med* 2015; 7 (299): 299ra122.
39. Nemati S, Holder A, Razmi F, *et al.* An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Crit Care Med* 2018; 46 (4): 547–53.
40. Singer M, Deutschman CS, Seymour CW, *et al.* The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016; 315 (8): 801–10.
41. Kashyap R, Singh TD, Rayes H, *et al.* Association of septic shock definitions and standardized mortality ratio in a contemporary cohort of critically ill patients. *J Crit Care* 2019; 50: 269–74.