Cost and mortality impact of an algorithm-driven sepsis prediction system

Jacob Calvert, Jana Hoffman, Christopher Barton, David Shimabukuro, Michael Ries, Uli Chettipally, Yaniv Kerem, Melissa Jay, Samson Mataraso & Ritankar Das

To cite this article: Jacob Calvert, Jana Hoffman, Christopher Barton, David Shimabukuro, Michael Ries, Uli Chettipally, Yaniv Kerem, Melissa Jay, Samson Mataraso & Ritankar Das (2017): Cost and mortality impact of an algorithm-driven sepsis prediction system, Journal of Medical Economics, DOI: 10.1080/13696998.2017.1307203

To link to this article: http://dx.doi.org/10.1080/13696998.2017.1307203
Cost and mortality impact of an algorithm-driven sepsis prediction system

Jacob Calverta, Jana Hoffmanb, Christopher Bartonb, David Shimabukuroc, Michael Riesd, Uli Chettipallyb,e, Yaniv Keremf, Melissa Jaya, Samson Mataraso, and Ritankar Das

aDascena Inc., Hayward, CA, USA; bDepartment of Emergency Medicine, University of California San Francisco, San Francisco, CA, USA; cDivision of Critical Care Medicine, Department of Anesthesia and Perioperative Care, University of California San Francisco, San Francisco, CA, USA; dAdult Critical Care, Advocate Health, Oak Brook, IL, USA; eDepartment of Emergency Medicine, Kaiser Permanente South San Francisco Medical Center, South San Francisco, CA, USA; fDepartment of Clinical Informatics, Stanford University School of Medicine, Stanford, CA, USA; gDepartment of Emergency Medicine, Kaiser Permanente Redwood City Medical Center, Redwood City, CA, USA; hDepartment of Bioengineering, University of California Berkeley, Berkeley, CA, USA

ABSTRACT

Aims: To compute the financial and mortality impact of InSight, an algorithm-driven biomarker, which forecasts the onset of sepsis with minimal use of electronic health record data.

Methods: This study compares InSight with existing sepsis screening tools and computes the differential life and cost savings associated with its use in the inpatient setting. To do so, mortality reduction is obtained from an increase in the number of sepsis cases correctly identified by InSight. Early sepsis detection by InSight is also associated with a reduction in length-of-stay, from which cost savings are directly computed.

Results: InSight identifies more true positive cases of severe sepsis, with fewer false alarms, than comparable methods. For an individual ICU with 50 beds, for example, it is determined that InSight annually saves 75 additional lives and reduces sepsis-related costs by $560,000.

Limitations: InSight performance results are derived from analysis of a single-center cohort. Mortality reduction results rely on a simplified use case, which fixes prediction times at 0, 1, and 2 h before sepsis onset, likely leading to under-estimates of lives saved. The corresponding cost reduction numbers are based on national averages for daily patient length-of-stay cost.

Conclusions: InSight has the potential to reduce sepsis-related deaths and to lead to substantial cost savings for healthcare facilities.

Introduction

Severe sepsis and septic shock are among the leading causes of death in the US1,2, and are responsible for ~375,000 deaths annually3,4. In addition to taking many lives, sepsis places tremendous financial strain on the healthcare system, costing over $20 billion per year nationally in diagnosis and treatment costs5. The widely accepted traditional definition of sepsis is the presence of Systemic Inflammatory Response Syndrome (SIRS)6, together with a known or suspected infection. However, this criterion is a notoriously non-specific indicator of sepsis risk. Severe sepsis is defined as sepsis, together with associated organ dysfunction7, and septic shock additionally includes refractory hypotension8. Each of these escalating conditions is associated with an increase in length of stay, patient mortality, and cost of treatment5,9.

Despite many attempts to define and identify sepsis, including the newly proposed Sepsis-3 definition7, sepsis diagnosis is particularly challenging because of the complexity and heterogeneity in both origin and clinical manifestation. The presence of infection cannot always be reliably determined9, and dysregulated host response to infection can be difficult to assess. However, early diagnosis of sepsis has been shown to be critical for effective medical intervention. Studies have shown that early diagnosis and treatment of sepsis, such as early goal-directed therapy (EGDT), can reduce adverse patient outcomes from severe sepsis and septic shock10-12. In spite of this research, sepsis detection methods have changed minimally in the last few decades.

Clinical decision support (CDS) systems are designed to assist clinicians in diagnosis, medication, and patient management. Computerized CDS systems may be especially valuable in assisting clinicians with complex or difficult diagnoses, such as sepsis13. The increasing availability of electronic health records (EHR) facilitates the development of computerized tools that attempt to identify sepsis through the analysis of these records14,15. Despite their wide availability, these tools often suffer from low specificity or sensitivity, involve additional manual data entry or complex text interpretation, require extensive laboratory results, or fail to predict which patients are at highest risk for developing severe sepsis or septic shock16,17. In this article, we discuss the key features and compute mortality reduction and financial impact of an algorithm-driven biomarker, InSight. It can detect and predict the onset of each of sepsis, severe sepsis,
and septic shock using patient vital signs automatically extracted from the EHR, with a high sensitivity and specificity, and provide these results through a computerized CDS system\textsuperscript{18–20}.

**Description of InSight**

*InSight* is the result of a machine learning workflow, which can predict the onset of sepsis, severe sepsis, or septic shock hours before onset, using only basic patient chart information routinely available in the EHR. To train *InSight* to predict sepsis, we first label retrospective inpatient stays based on whether or not the patient developed sepsis, according to criteria collectively referred to as a ‘gold standard’. We then choose a set of clinical measurements, the values, correlations, and trends of which (collectively called “features”) are associated with the gold standard labels of septic or non-septic. When *InSight* is used clinically, these measurements are autonomously collected and used to generate predictions of sepsis. *InSight*'s prediction takes the form of a numerical score, which indicates the likelihood of a given patient having sepsis or developing sepsis in the near future.

*InSight* is flexible with regard to what clinical data it uses to make features, and it can be adapted to site-specific parameters. In our prior work, we have used clinical measurements which are nearly universally available at the bedside, such as heart rate, respiration rate, temperature, systolic blood pressure, pulse pressure, peripheral capillary oxygen saturation, Glasgow Coma Score, and age. *InSight* can make reliable predictions, despite the simplicity of the clinical measurements it uses, because it analyzes vital signs over time to extract trend information, as well as the correlations between pairs and higher order correlations of patient measurements. These features provide more refined knowledge of the patient condition, because complex conditions such as sepsis often impact the feedback mechanisms regulating the normal correlations of vital signs. Notably, *InSight* does not require the use of waveform data, laboratory results, or the complicated interpretation of free text notes. There are substantial advantages to analyzing only basic clinical measurements, which include:

1. **Prevalence**—these measurements are commonly available for most patients at the bedside.
2. **Integration**—vital signs are routinely recorded in a hospital’s EHR as a part of normal clinical practice, and autonomously accessing this information does not interrupt workflow.
3. **Relevance**—common vital signs, such as temperature and respiration rate, are highly relevant to the development of sepsis.
4. **Frequency**—most vital sign measurements are sampled frequently in inpatient and emergency settings, providing timely updates on the patient’s status.

In addition to the freedom clinicians and hospital administrators have in choosing the clinical measurements given as input to *InSight*, there is flexibility in the choice of gold standard for identifying sepsis patients. Because the gold standard shapes *InSight*'s training and the predictions it ultimately makes, *InSight* can be adapted by retraining it with a different sepsis standard. Furthermore, this gold standard can be modified to train an *InSight*-like algorithm-driven biomarker to predict patient conditions other than sepsis, such as in-hospital mortality\textsuperscript{19} or patient stability\textsuperscript{20}.

*InSight* has been validated through several studies in which it was used to predict sepsis onset for three different gold standards: SIRS criteria\textsuperscript{18}, septic shock\textsuperscript{21}, and sepsis (as per the recent redefinition of Sepsis-3, roughly equivalent to severe sepsis as per Sepsis-2)\textsuperscript{22}. Table 1 summarizes several relevant metrics for the detection of severe sepsis, illustrating that *InSight* has outperformed commonly used existing disease severity scoring systems such as SIRS, MEWS (Modified Early Warning Score)\textsuperscript{23}, SAPS II (Simplified Acute Physiology Score)\textsuperscript{24}, or SOFA (Sequential Organ Failure Assessment)\textsuperscript{25}. These metrics will be used in the next section, where we compute the expected cost and mortality reductions associated with using *InSight* in lieu of an existing scoring system for a variety of clinical settings.

**Methods**

We place our cost and mortality estimations in the context of an intensive care unit (ICU) which is already using an early warning score (EWS) for severe sepsis. For illustration, we compare with two severe sepsis screening risk scores, SOFA and SIRS. For the purposes of our calculations, we consider that clinicians refer to *InSight* predictions at the same time as the predictions from SOFA and SIRS. Furthermore, the score cut-off thresholds for *InSight*, SOFA, and SIRS are fixed such that they have similar specificities (i.e. holding the false alarm rate constant), and, thus, isolate the benefit of using *InSight* to an improvement in prediction sensitivity.

We consider predictions made 0, 1, and 2 h early, which we term “lookaheads” of 0, 1, and 2 h. At equal specificities, we tabulate the differences in sensitivity, \(\Delta_{s}\), and perform a calculation according to the procedure below. Note that this process is done for both SOFA and SIRS, but only illustrated below for SOFA.

For a “lookahead” of \(k\)-hours, there is a difference in sensitivity between *InSight* and SOFA of \(\Delta_{s}\). The product of the difference in sensitivity, and the annual number of sepsis patients seen by an ICU, \(N\), gives the increase in true positives found by *InSight* over SOFA. For ease of presentation,
we split the $N$ term (number of sepsis encounters) into the prevalence of sepsis, $p$, and the total number of patients, $n$.

Increase in true positives = $\Delta_k * p * n$

The median time-to-treatment for sepsis has been reported to be 6 h$^{12}$. We assume that, by producing an alert $k$-hours early, clinicians are able to respond in $(6 - k)$ hours. Ferrer et al.$^{26}$ have investigated the effect of delayed antibiotic administration on mortality for cases of both severe sepsis and septic shock, using generalized estimating equation population averaged logistic regression in a multi-center retrospective analysis. They have found that administration of antibiotic within the first hour of severe sepsis or septic shock onset leads to a survival probability of 75.4%. However, each hour of delay in antibiotic administration decreases survival by 1.42% per hour. Therefore, we compute that the additional true positive cases identified by InSight have a survival rate of $0.754 - 0.0142 * (5 - k)$. This description leads to the calculation of InSight saving an additional number of patients equal to

$$\Delta_k * (0.754 - 0.0142 * (5 - k)) * p * n$$

To estimate the cost reduction, we assume that all septic patients are eventually treated for sepsis at a cost that is independent of their early identification by InSight, SOFA, or SIRS. In this way, we attribute all cost reductions to a length-of-stay (LOS) reduction. Ideally, we could sum the contributions to mortality and cost reduction, resulting from patients identified at onset, as well as those detected early (i.e. summing over “lookaheads”, $k$). Given that the difficulty of identifying a septic patient at onset is correlated with the difficulty of identifying them early, we cannot sum the mortality and cost reductions from each hour of lookahead. In order to be conservative in our calculations, we use only the number of additional true positives identified at onset. We assume that an identification at onset improves compliance with 6-h sepsis bundles, which reduces LOS by an average of 3.7 days$^{27}$. With the daily average cost of a patient stay of $2,000 per day$^{28}$, this LOS reduction amounts to $7,400 per additional true positive.

Based on our earlier analysis, we calculate that $7,400 * \Delta_k^* (0.754 - 0.0142 * (5 - k)) * p * n$ is saved each year for the $p * n$ sepsis cases that year. To get a sense for the size of this number, note that there are up to 375,000 sepsis-related deaths each year$^{3}$, out of more than 5.7 million ICU visits$^{29}$. With $\Delta_k = 0.10$, $k = 4$, $n = 5.7$ million, and $p = 0.126^{30}$, we calculate that nationwide adoption of InSight will save more than 53,000 lives, or $\sim$14% of all sepsis-related deaths, in the ICU alone. The total reduction in cost attributed to shorter patient stays is $\sim$ $393 million.

To assess the impact of using InSight for individual hospitals, we take our population-level cost and mortality reduction results and contextualize them in terms of a cost and mortality reduction per year, for a given number of beds. For example, consider that, in 2000, there were 67,579 ICU beds in the US$^{31}$. Around the same time, there were $\sim$4 million ICU admissions each year$^{22}$. To model an “average” hospital, we uniformly spread these admissions across the ICU beds. We conclude that there are $\sim$60 ICU admissions per ICU bed per year, with an average length of stay of 6 days. So, for a hospital with 100 ICU beds, we can set $n = 6,000$ patients and apply the formulae above.

**Results**

We now quantify the financial and medical advantages of using InSight over current sepsis screening systems with the methodology presented above. Based on receiver operating characteristic curves computed using the recent Sepsis-3 definition$^{7}$, we determine the sensitivities and specificities for InSight, SOFA, and SIRS at various thresholds. To assess the number of additional lives saved using InSight, we fix specificity and calculate the change in sensitivity, $\Delta_k$, when compared to SIRS and SOFA. When fixing the specificity at a value between 0.80–0.90, InSight demonstrates a significant improvement in the number of true positive sepsis cases caught at onset and in advance (Figure 1a). We determine the number of additional cases caught by multiplying $\Delta_k$ by the number of septic patients. InSight also reduces

![Figure 1](image-url)
false alarms. Alarm fatigue can be detrimental in hospitals, leading to the potential for dismissal of real cases. By fixing sensitivity at 0.70, we calculate the change in specificity when using InSight over other scoring systems. We multiply this difference by the number of non-septic patients to determine the reduction in false alarms that clinicians will experience. Figure 1(b) quantifies the false alarm reduction that InSight provides over SIRS and SOFA.

InSight produces fewer overall alarms per day, while still maintaining a higher sensitivity and specificity as compared with SOFA and SIRS. For example, let us consider 100 beds and a sepsis prevalence of 12.6%. As estimated earlier, with ~60 patients per bed per year, each patient spends an average length of stay of 6 days. To simplify calculations, for septic patients, we assume onset occurs on a randomly selected day during a patient’s 6-day stay. We, therefore, expect 12.6 of these patients to have sepsis by the end of the stay, and 2.1 of these patients to present with sepsis each day. InSight will trigger (2.1*sensitivity) alarms corresponding to true positives and (97.9*(1 – specificity)) alarms corresponding to false positives. Figure 2 displays the number of alarms InSight, SOFA, and SIRS trigger per day, depending on the sensitivity chosen. At a sensitivity of 0.80, InSight generates 46 fewer alarms than SIRS per day.

With a reduction in false alarms, and an increase in true positive rate, InSight can both reduce sepsis-related hospital costs and save more lives than traditional approaches. Table 2 summarizes InSight’s calculated financial and health impacts compared with SOFA and SIRS.

**Discussion and conclusion**

Early identification of patients at high risk for developing life-threatening sepsis syndromes is critical for providing effective medical care and interventions. Various disease severity scoring systems (such as MEWS, SAPS II, and SOFA) are frequently employed in an attempt to identify patients most in need of medical resources, but often lack the desired sensitivity, resulting in increased numbers of false alerts. With the advances in electronic health records, the opportunity for continuous and accurate monitoring of patient condition is becoming more achievable, and provides the opportunity to save lives and reduce costs associated with difficult-to-detect medical conditions such as severe sepsis. InSight uses data commonly entered in the EHR, analyzes the correlations between common patient vital signs, and generates accurate sepsis alerts and predictions which outperform current screening systems. To translate the increase in sensitivity and specificity provided by InSight into concrete measures, we have created a simple, conservative model for computing the financial and mortality savings as a function of the number of beds, arising from the adoption of InSight for sepsis detection.

Our model demonstrates that InSight significantly increases the number of septic patients that can be identified over the number of patients identified by SOFA and SIRS, and also provides a sizable reduction in the number of false alarms. This increased accuracy of sepsis identification, combined with timely sepsis bundle compliance, leads to annual cost savings in the millions of dollars and hundreds of lives saved per year in large hospital settings. The increase in accuracy of InSight over these scores is particularly notable, because InSight utilizes only vital sign data, which is readily obtained at the bedside. In contrast, scores such as SOFA and SAPS II rely on laboratory tests such as white blood cell counts, immature band cell determination, renal and liver function panels, metabolic tests, and prior medical history to

![Figure 2. Number of alarms per day using InSight, SOFA, and SIRS for sepsis detection with 100 beds (a) and 300 beds (b).](image)

**Table 2.** InSight’s advantage in lives saved and cost savings annually over SOFA and SIRS identification of septic patients, as a function of the number of beds. Based on prior performance, a facility switching to InSight from SIRS or from SOFA can expect to save an average of $925 per bed per month or $574 per bed per month, respectively. Similarly, upon switching to InSight, ~1–1.5 additional patients per bed per year are expected to survive.

<table>
<thead>
<tr>
<th>Number of beds</th>
<th>50</th>
<th>100</th>
<th>300</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>InSight vs SOFA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives saved (/year)</td>
<td>47</td>
<td>93</td>
<td>279</td>
<td>466</td>
</tr>
<tr>
<td>Cost savings ($100k/year)</td>
<td>3.4</td>
<td>6.9</td>
<td>21</td>
<td>34</td>
</tr>
<tr>
<td><strong>InSight vs SIRS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives saved (/year)</td>
<td>75</td>
<td>150</td>
<td>451</td>
<td>752</td>
</tr>
<tr>
<td>Cost savings ($100k/year)</td>
<td>5.6</td>
<td>11</td>
<td>33</td>
<td>55</td>
</tr>
</tbody>
</table>

J. CALVERT ET AL.
generate scores. The machine learning approach of InSight leads to higher sensitivity and specificity (AUROC) and DOR for severe sepsis than these scores, without requiring laboratory results, as shown in Table 1. The minimal input requirements of InSight provide a potential clinical advantage in ease of implementation and rapid diagnosis.

There are a few aspects of our financial estimation which invite discussion, due in large part to its simplicity. In order to compute cost and mortality savings most conservatively, we have proposed a use of InSight which differs from its mode of clinical action. Rather than InSight generating necessary alerts for the immediate attention of the clinician, for the purposes of this calculation, we limit InSight to only activate at the same point in time as SIRS and SOFA. InSight demonstrates maximal sensitivity and specificity at the time of sepsis onset, and also outperforms SIRS and SOFA prior to onset. This comparison illustrates InSight’s increased true positive rate and true negative rate over commonly used scoring systems when run at discrete points in time. In other words, the medical and health implications considered here are conservatively derived from a single fixed-time, rather than continuous-time use, which would more accurately reflect InSight’s actual use case, and is likely to result in higher cost savings and increased mortality reduction than reported here.

A second limitation of our model concerns its generalizability to various hospital environments. Using averaged national data, we also assume the number of ICU admissions and beds is uniform across hospitals. Thus, we determine that each patient spends an average of 6 days in the ICU, which is likely to vary across hospitals. Tailoring these averages to specific ICUs may alter the cost savings and alarm counts presented in this report. Furthermore, treatment costs differ by hospital and region. We estimate the cost of each patient-day to be ~ $2,000, without adjusting for regional or medical variability. While patients treated for sepsis may cost more than other inpatients, we also assume that patient costs per day are averaged across all patient stays, regardless of patient medical condition, to keep our calculations conservative.

Our assumptions surrounding standard bundle compliance is an additional point of discussion. The model uses a few reasonable metrics of outcomes for sepsis and early intervention based on published data, including the median time to sepsis intervention, discussed in the Ferrer et al. paper. By using this value, we assume that clinicians will improve bundle compliance when adopting a sepsis screening tool. Time to bundle initiation varies widely, based on hospital and clinician training. We expect that, by using an automated screening system, clinicians will adhere to 6-h bundle requirements and significantly improve compliance. The rate at which hospitals actually comply with and initiate the sepsis bundle after sepsis identification will impact the financial savings and number of lives saved using InSight.

Despite several limitations, our simple methodology clearly exhibits InSight’s impact on mortality and cost reduction. By accurately predicting which patients are most likely to become septic, clinicians are able to deploy sepsis bundles earlier, which leads to better outcomes for patients. A natural consequence of more effective and rapid treatment is a reduction in LOS, which directly translates as cost savings. Thus, while improved outcomes often come at a higher cost in medicine, the opportunity of initiating earlier sepsis care by using an algorithm-driven biomarker is one where cost reduction and quality improvement incentives align. Ultimately, healthcare systems, providers, and patients all benefit.

Transparency statement

Declaration of funding

This material is based upon work supported by the National Science Foundation under Grant No. 1549867. The funder had no role in the conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, and approval of the manuscript; and decision to submit the manuscript for publication.

Declaration of financial/other interests

All authors who have affiliations listed with Dascena (Hayward, CA) are employees of Dascena. Dr. Barton reports receiving consulting fees and grant funding from Dascena. Dr. Shimabukuro reports receiving grant funding from Dascena. JME peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Acknowledgments

We gratefully acknowledge Thomas Desautels, PhD, for the severe sepsis InSight experiments and results, and we acknowledge the assistance of Nima Shahjarian in the preparation of this manuscript. We acknowledge Dr Grant Fletcher for helpful comments and editing assistance, with thanks. We acknowledge Qingqing Mao and Hamid Mohamadlou for significant contributions to the development and application of the machine learning algorithm, InSight.

ORCID

Jacob Calvert http://orcid.org/0000-0001-9173-0946
Jana Hoffman http://orcid.org/0000-0002-7745-3900
Christopher Barton http://orcid.org/0000-0002-5842-2949
David Shimabukuro http://orcid.org/0000-0002-6742-9173
Uli Chettipally http://orcid.org/0000-0001-8833-5478
Yaniv Kerem http://orcid.org/0000-0002-6855-710X
Melissa Jay http://orcid.org/0000-0002-8740-4803
Samson Mataraso http://orcid.org/0000-0003-3146-2243
Ritankar Das http://orcid.org/0000-0002-1326-844X

References