Delay Within the 3-Hour Surviving Sepsis Campaign Guideline on Mortality for Patients With Severe Sepsis and Septic Shock

Lisiane Pruinelli, PhD, RN; Bonnie L. Westra, PhD, RN, FAAN, FACMI1,2; Pranjul Yadav, PhD3; Alexander Hoff3; Michael Steinbach, PhD3; Vipin Kumar, PhD3; Connie W. Delaney, PhD, RN, FAAN, FACMI1,2; Gyorgy Simon, PhD2,4

Objectives: To specify when delays of specific 3-hour bundle Surviving Sepsis Campaign guideline recommendations applied to severe sepsis or septic shock become harmful and impact mortality.

Design: Retrospective cohort study.

Setting: One health system composed of six hospitals and 45 clinics in a Midwest state from January 01, 2011, to July 31, 2015.

Patients: All adult patients hospitalized with billing diagnosis of severe sepsis or septic shock.

Interventions: Four 3-hour Surviving Sepsis Campaign guideline recommendations: 1) obtain blood culture before antibiotics, 2) obtain lactate level, 3) administer broad-spectrum antibiotics, and 4) administer 30 mL/kg of crystalloid fluid for hypotension (defined as “mean arterial pressure” < 65) or lactate (> 4).

Measurements and Main Results: To determine the effect of t minutes of delay in carrying out each intervention, propensity score matching of “baseline” characteristics compensated for differences in health status. The average treatment effect in the treated computed as the average difference in outcomes between those treated after shorter versus longer delay. To estimate the uncertainty associated with the average treatment effect in the treated metric and to construct 95% CIs, bootstrap estimation with 1,000 replications was performed. From 5,072 patients with severe sepsis or septic shock, 1,412 (27.8%) had in-hospital mortality. The majority of patients had the four 3-hour bundle recommendations initiated within 3 hours. The statistically significant time in minutes after which a delay increased the risk of death for each recommendation was as follows: lactate, 20.0 minutes; blood culture, 50.0 minutes; crystalloids, 100.0 minutes; and antibiotic therapy, 125.0 minutes.

Conclusions: The guideline recommendations showed that shorter delays indicate better outcomes. There was no evidence that 3 hours is safe; even very short delays adversely impact outcomes. Findings demonstrated a new approach to incorporate time t when analyzing the impact on outcomes and provide new evidence for clinical practice and research. (Crit Care Med 2018; XX:00–00)

Key Words: electronic health record data; mortality; predictive modeling; severe sepsis; surviving sepsis campaign guideline

Sepsis is a major health problem with increasing prevalence, high costs, and poor outcomes. In the United States, sepsis accounts for about 5.2% of hospital expenditures, more than $20 billion per year (1) and increased hospitalizations for more than 1 million people in 2008 (2). Between 1999 and 2014, there were 2,470,666 deaths with sepsis on the death certificate, and annual sepsis-related deaths increased by 31% (2, 3). Severe sepsis and septic shock are a subset of sepsis conditions with at least one organ dysfunction (2). This subset of conditions accounts for 25–50% of in-hospital mortality (2), ranging from 45% to 65% after 6 months of discharge. If patients with these
conditions survive, they have increasing chronic complications, morbidity; high costs of care, and decreasing quality of life (4, 5). Due to the seriousness of these outcomes, the focus of this study is on severe sepsis and septic shock.

The Surviving Sepsis Campaign (SSC) international guidelines (6) were developed for early detection and treatment of severe sepsis and septic shock. Although there is agreement that applying the SSC guidelines improves outcomes, there are several challenges including varying strength of evidence associated with the recommendations (6). Since 2001, the definitions and guidelines for sepsis have continuously evolved (6). In 2004, there were 46 recommendations and pediatric management considerations which were subsequently divided into 6- and 24-hour bundles (7). The guidelines were revised in 2008, 2012 (6), and most recently updated in 2015 (8). The recommendations that repeatedly proved effective were simplified and grouped into 3- and 6-hour bundles. The next question is whether there is more specific timing that would improve outcomes. The 3-hour recommendations, which must be carried out within 3 hours from the first time sepsis is suspected, are as follows: 1) obtain a blood culture before antibiotics, 2) obtain a lactate level, 3) administer broad-spectrum antibiotics, and 4) administer 30 ml/kg of crystalloid fluid for hypotension (defined as a “mean arterial pressure” [MAP] < 65) or lactate (> 4).

Previous studies examined the effect of the SSC guideline on various outcomes, such as hospital and ICU length of stays, discharge disposition, hospital readmission, and mortality (5, 9–12). There are mixed results regarding a decrease in mortality, which is likely related to outcome definitions, types of interventions evaluated, population heterogeneity, and timing (13–21). It is unclear if the timing of specific SSC guideline recommendations within the 3-hour bundle improves the probability of survival. The availability of electronic health record (EHR) data and newer machine learning analytics may provide new insights for disease management and clinical improvement. Sepsis studies using EHR data have focused on trends in sepsis (22) and early identification of sepsis and/or the impact on outcomes (23–27). A few studies used machine learning, a family of techniques that use artificial intelligence to learn from large collections of data, to develop predictive models for sepsis outcomes (28–31). None focused on evaluating the timing of SSC guideline recommendations within the 3-hour bundle to determine if a delay in these recommendations results in harm. We previously used machine learning and found the SSC guidelines to be effective on outcomes of mortality and complications (32); however, we did not look at the timing within the 3-hour bundle to determine if there is a harmful effect in less than 3 hours. The purpose of this study was to specify the time frame when the delay of specific 3-hour bundle guideline recommendations applied to severe sepsis or septic shock becomes harmful and impacts mortality.

MATERIALS AND METHODS

Study Design and Data Selection

A retrospective cohort study was conducted using data from all adult patients who were hospitalized with a billing diagnosis of severe sepsis or septic shock between January 01, 2011, and July 31, 2015. After institutional review board approval (number 1311E46042), a deidentified dataset was obtained from a clinical data repository which contained over 2.4 million patients. The inclusion criteria were patients 18 years old and older, both medical and surgical patients, who were hospitalized with a billing diagnosis of severe sepsis or septic shock (International Classification of Diseases, 9th Edition codes 995.92 and 785.5* excluding 785.51). The initial dataset contained 5,374 patients. The first hospitalization was selected as the index encounter. Exclusion criteria were as follows: did not meet two of the criteria for sepsis suspicion (n = 114) or no data for antibiotic therapy (n = 142), respiratory rate (RR) (n = 3), or WBC count (n = 43). This resulted in a final sample of 5,072 patients. The onset of sepsis (baseline) was estimated as the earliest date and time when the patient met at least two of the following six criteria: MAP less than 65, heart rate (HR) greater than 100, RR greater than 20, temperature less than 95 or greater than 100.94, WBC count less than 4 or greater than 12, and lactate greater than 2.0 (33).

Variables of Interest

Variables of interest included predictors, “delay” in administering guideline recommendations, and outcomes. The 16 predictor variables were age, gender, race, ethnicity, lactate, WBC count, HR, RR, temperature, MAP, Charlson Comorbidity Index, and five Comorbidity Severity Scores, which describe the severity of preexisting diseases in each of the following five health categories: respiratory, cardiac, liver, kidney, and vascular system. Using domain experts, Comorbidity Severity Scores were computed as the weighted sum of the diagnoses, identified from the problem list and billing diagnoses prior to the index encounter, where the weights were assigned to reflect severity and acuteness: 0 = chronic, 1 = mild, 2 = moderate to severe. These weights were summed for each Comorbidity Severity Score and each patient for analysis, resulting in a continuous scale. “Delay” of obtaining a lactate level and blood culture or administering a broad-spectrum antibiotic was measured from baseline to the time these actions occurred (up to 6hr); for the recommendation of “administer crystalloids when lactate is greater than 4 mmol/L,” “delay” is the time between discovering that lactate is high (time of the laboratory result) and the time of administering crystalloids. The quantity of interest is the average causal effect of “delay” on all-cause in-hospital mortality. The outcome is mortality occurring at any time during that hospital stay.

Data Analysis

The purpose of the analysis was to determine the effect of t minutes of delay in carrying out a recommendation, for any amount of delay t between 15 and 360 minutes. This effect was defined as the average treatment effect in the treated (ATT) and is interpreted as the reduction in the probability of in-hospital mortality as a result of carrying out a particular guideline recommendation with less than t minutes of delay. We performed sequential propensity score matching (PSM) at t = 15, 30, ..., 360 minutes with the “exposed” group consisting of patients...
who received a particular guideline recommendation in less than $t$ minutes, versus the “control” group who received the recommendation with more than $t$ minutes of delay. Patients who did not receive the guideline-recommended treatment before discharge were excluded. The sequential PSM procedure resulted in 24-point estimates of ATT at 24 discrete values of delay. We applied a local polynomial smoother (loess) to the 24-point estimates to obtain a smooth ATT curve from which the effect of delay can be determined for any $t$ between 15 minutes and 6 hours in a nonparametric fashion. The same procedure was repeated for all four guideline recommendations.

The predictors of the propensity score model included predictors of the delay as well as predictors of the outcome. These were age, sex, the baseline laboratory results and vitals, and the severity score of the preexisting comorbidities. The MatchIt R package (version 3.0.1) was used for propensity matching, using nearest neighbor matching with a caliper of 0.25 SDs. We used the paired $t$ test as well as the standardized mean difference (SMD) between the exposed and control groups to assess the achieved balance (34). A $p$ value of 0.05 and a SMD of less than 0.2 were required to achieve significance. All 96 propensity-matched populations (24 for each of the four recommendations) had a SMD of less than 0.2, but some matched populations (for Crystalloid beyond 3-hr delay) failed the $t$ test. Furthermore, this did not impact the ATT because the SMD was still less than 0.2. In the matched population, ATT was computed as the difference in the proportions of patients who died in the hospital between the exposed and control groups.

The 95% CI for the ATT curve was obtained through bootstrap simulation with 1,000 replications. Continuous variables are reported as mean and interquartile range and categorical variables as a count and percentage. To determine statistical significance, a 5% empirical significance level was adopted.

**RESULTS**

There were 5,072 patients with severe sepsis or septic shock during the study period. Table 1 describes the baseline characteristics of the sample and variables used for each PSM. There were 1,412 patients (27.8%) who died in the hospital. The majority of patients had the four 3-hour bundle recommendations initiated within 3 hours. The percentage of patients who did not receive the recommendation within 3 hours included lactate (14.3%), blood culture collected before antibiotic therapy (8.2%), broad-spectrum antibiotics initiated (2.7%), and crystalloids initiated inside the recommended 3 hours (0.3%). From these 214 patients (4.2%) who did not receive any of the above recommended actions within 3 hours, 88 (41.1%) died.

Figure 1A illustrates the usual effect of delays in blood culture draw on in-hospital mortality. The horizontal axis corresponds to “delay” in minutes, and the vertical axis corresponds to the ATT. The solid line is the (smoothed) ATT at each time point, and the dashed lines correspond to the 95% CI. The arrow indicates the $t$ minutes delay when the recommendation becomes statistically significant. For example, the ATT at 30 minutes is 0.006 (95% CI, –0.026 to 0.034), meaning that a delay of half an hour increases the risk of mortality by 0.006 (six additional mortalities per 1,000 patients), but it is not statistically significant. A delay of 50 minutes, however, increases the patient’s risk of mortality by 0.028 (CI, 0.004–0.058) suggesting that a delay of 50 minutes significantly increases the risk of mortality. The distance between these two lines is the effect of the delay and is the solid line in Figure 1A. The solid line in Figure 1B shows the survival probability for treated (patients with less than $t$ minutes of delay in drawing a blood culture) and the dashed line shows for the control patients (those with $t$ or greater minutes of delay). Figure 1B helps gauge how big the effect of the delay is relative to the actual survival probability. Similarly, Figures 2-4 illustrate results for the lactate, antibiotic therapy, and crystalloids recommendations, respectively.

**DISCUSSION**

The effect of “delay” for each SSC 3-hour bundle recommendation on in-hospital mortality for patients with severe sepsis and septic shock was evaluated. A nonparametric methodology of sequential PSM for calculating the ATT at a sequence of time points was used to estimate the effect of implementation delay in the SSC guideline recommendations on all-cause in-hospital mortality. We found that delays in administering all four guideline recommendations, even when they did not

---

**TABLE 1. Predictor Variables for Patients With Severe Sepsis and Septic Shock (n = 5,072)**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Median (IQR or n %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>66 (55–78)</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>2,588 (51.0)</td>
</tr>
<tr>
<td>Race (Caucasian)</td>
<td>4,349 (85.7)</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic)</td>
<td>4,822 (95.0)</td>
</tr>
<tr>
<td>Lactate</td>
<td>1.8 (1.8–3.2)</td>
</tr>
<tr>
<td>WBCs</td>
<td>12.6 (7.9–17.4)</td>
</tr>
<tr>
<td>Median blood pressure</td>
<td>81.3 (67–96)</td>
</tr>
<tr>
<td>Heart rate</td>
<td>97 (90–103)</td>
</tr>
<tr>
<td>Temperature</td>
<td>98.5 (97–100)</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>20 (18–24)</td>
</tr>
<tr>
<td>Charlson Comorbidity Index</td>
<td>3.9 (1–6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comorbidities Severity Scores</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory</td>
<td>2.2 (0–4)</td>
</tr>
<tr>
<td>Cardiac</td>
<td>5.6 (1–9)</td>
</tr>
<tr>
<td>Liver</td>
<td>0.5 (0–0)</td>
</tr>
<tr>
<td>Cerebrovascular</td>
<td>0.3 (0–0)</td>
</tr>
<tr>
<td>Kidney</td>
<td>1.3 (0–2)</td>
</tr>
</tbody>
</table>

IQR = interquartile range.
exceed 3 hours, were associated with a significant increase in in-hospital mortality.

Overall, this study showed that no delay is safe, although modest delays may not be harmful in a clinically meaningful way. For most outcomes, the ATT curves (solid curve in Figs. 1A, 2A, 3A, and 4A) are uniformly above zero and are strictly monotonically increasing. The longer the delay, the more the harm. Up to a certain delay, the effect is not statistically significant. This does not mean that this much delay is not harmful; it simply means that we do not have sufficient evidence to show the (potential) harm. Statistical significance depends on sample sizes, and the delay that is associated with statistically significant harm could be different across hospitals. Thus, the recommended actions can and should be carried out as soon as possible to prevent harm.

The shape of the curves provides further support for our recommendation. The curves indicate that harm accumulates very quickly in the beginning (short delays) but slows down for longer delays and flattens out by about 3 hours. This means that delays exceeding the 3 hours are associated with little additional harm on top of the harm the patient has experienced within 3 hours. We acknowledge that the high compliance rate with the 3-hour bundle may influence our results: the minimal additional harm after 3 hours of delay may be partly artefactual, attributable in part to the fact that most patients have already received the recommended intervention within 3 hours, and the sample size is too small beyond 3 hours of delay to determine further harm. However, if the observed effect was indeed artefactual, we would see the CI widen substantially (as it does for crystalloid use) and also propensity matching would fail. During our analysis, we always observed sufficient support (overlap between the treatment and control groups) for propensity matching and adequate balance of the variables (except crystalloid beyond 3 hr).
Results showed an in-hospital mortality of 27.8%. This is consistent with other reports that showed similar rates worldwide, specifically in the United States and Europe, although rates may vary greatly according to different geographic regions (17, 21). Variability in outcomes may be due to the high heterogeneity of sepsis populations; this study considered incorporating a large variety of health conditions into account when performing PSM. Furthermore, diagnosis differences and, consequently, earlier initiation of treatment vary greatly as a result of several human, institutional, and financial practices (21). Previous studies reported evidence that timing of antimicrobial therapy in the presence of septic shock is associated with higher mortality (6, 8). Our findings agree that mortality increases with delay in antibiotics, but only becomes significant at nearly 3 hours after baseline (8).

There are several strengths in this study. The adopted methodology of sequential PSM in conjunction with bootstraps is novel and necessary to minimize the confounding that occurs in observational studies. We were able to incorporate several covariates in our model, making the matched pairs very alike. This combined methodology provided the estimated effect of treatment considering matched pairs with similar propensity score values, thus, decreasing bias estimation. Another strength of the study is the use of EHR data. While results from EHR-based observational studies are often considered preliminary that requires validation through clinical trials, studying the effect of delaying treatment for research purpose would be unethical, and thus designing a trial for this purpose would be difficult.

Limitations exist in this study, such as the reuse of EHR for research from one single clinical source of data. Our dataset only provided in-hospital mortality, and we were not able to take into account mortality that occurred outside the hospital or long-term outcomes. We acknowledge that we did not adjust for baseline vasopressors and mechanical ventilation use, and both are important predictors of mortality among sepsis patients. Another limitation is the identification of a consistent list of broad-spectrum antibiotics. Our approach may have overestimated this, as a single list of IV broad-spectrum antibiotics was not found using national data standards nor in the current literature or consulting with two pharmacists. Although our efforts to address this gap, a clear listing of IV broad-spectrum antibiotics is still needed for comparison across studies, particularly in phenotyping EHR data and mapping to sepsis guideline recommendations (8, 35). Likewise, the compliance with crystalloid recommendation could be overestimated, as this study was not able to extract precise data that could account for volume administered and/or the type of crystalloids. Future efforts are needed to standardize nonstructured data, such as flowsheet data where crystalloids volume is documented; the researchers could capture and reuse this type of information for comparative effectiveness research across organizations (34). The definition of “delay” does not depend on the length of exposure to sepsis. The actual onset time of sepsis is unknown; it could have occurred outside the hospital. While using actual onset time reflects physiology better, using “first suspicion” reflects clinical reality better: patient’s lactate cannot be measured before they come to the hospital.

CONCLUSIONS

The guideline recommendations should be applied as soon as sepsis is suspected; applying the guidelines with shorter delays is associated with better outcomes. Our work found no evidence that 3 hours is safe; we found evidence that even very short delays can adversely impact outcomes. Our study demonstrated a new approach to incorporate time (“delay”) when analyzing the impact on outcomes and provide new evidence for clinical practice and research. It adopted a novel methodology using EHR data and bootstrapping in conjunction with PSM to estimate the ATT of SSC guideline recommendations “delay” on in-hospital mortality. The use of EHR data is important because a clinical trial that delays urgent treatment would be unethical. Based on these findings, future studies should also investigate the impact on long-term outcomes, including national or international dataset of clinical data to account for geographic variability and heterogeneity.

REFERENCES


