NEW MANDATED CENTERS FOR MEDICARE AND MEDICAID SERVICES
REQUIREMENTS FOR SEPSIS REPORTING: CAUTION FROM THE FIELD

Emily L. Aaronson, MD,*† Michael R. Filbin, MD,* David F. M. Brown, MD,* Kathy Tobin, RN,† and
Elizabeth A. Mort, MD, MPH†‡

*Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, †Lawrence Center for Quality and Safety, Massachusetts General Hospital and Massachusetts General Physicians’ Organization, Boston, Massachusetts, and ‡Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts

Corresponding Address: Emily L. Aaronson, MD, Department of Emergency Medicine, 0 Emerson Place, Suite 3B, Boston, MA 02115

Abstract—Background: The release of the Center for Medicare and Medicaid Service’s (CMS) latest quality measure, Severe Sepsis/Septic Shock Early Management Bundle (SEP-1), has intensified the long-standing debate over optimal care for severe sepsis and septic shock. Although the last decade of research has demonstrated the importance of comprehensive bundled care in conjunction with compliance mechanisms to reduce patient mortality, it is not clear that SEP-1 achieves this aim. The heterogeneous and often cryptic presentation of severe sepsis and septic shock, along with the multifaceted criteria for the definition of this clinical syndrome, pose a particular challenge for fitting requirements to this disease, and implementation could have unintended consequences. Objective: Following a simulated reporting exercise, in which 50 charts underwent expert review, we aimed to detail the challenges of, and offer suggestions on how to rethink, measuring performance in severe sepsis and septic shock care. Discussion: There were several challenges associated with the design and implementation of this measure. The ambiguous definition of severe sepsis and septic shock, prescriptive fluid volume requirements, rigid reassessment, and complex abstraction logic all raise significant concern. Conclusions: Although SEP-1 represents an important first step in requiring hospitals to improve outcomes for patients with severe sepsis and septic shock, the current approach must be revisited. The volume and complexity of the currently required SEP-1 reporting elements deserve serious consideration and revision before they are used as measures of accountability and tied to reimbursement. © 2016 Elsevier Inc. All rights reserved.

Keywords—sepsis; health care quality; access and evaluation; quality indicators; health care quality assessment

INTRODUCTION

The release by the Center for Medicare and Medicaid Services (CMS) of its latest quality measure, the Severe Sepsis/Septic Shock Early Management Bundle (SEP-1), has intensified the long-standing debate regarding the optimal early diagnosis and management of the patient with severe sepsis or septic shock. Before the initiation of the first Surviving Sepsis Campaign, severe sepsis and septic shock were responsible for approximately one in three hospital deaths, and carried with them a mortality rate ranging from 35% to 60% (1–3). The costs of treating these patients is also substantial, with mean hospital reimbursements estimated at greater than $29,000 per case (4–6). The cost to the United States (US) health care system during the previous decade has been estimated to exceed $16.7 billion annually (6). More recent estimates suggest that the annual costs for sepsis care exceed $20 billion, and are increasing more than 10% each year (7).
The last decade of research has demonstrated the importance of comprehensive bundled care in conjunction with compliance mechanisms to reduce patient mortality (8). Although current evidence indicates that standardized order sets and comprehensive quality-improvement infrastructure are effective mechanisms to promote early identification and prompt therapy for patients with severe sepsis and septic shock and to positively influence physicians’ clinical decision making, significant delays in care are still commonplace for these conditions and compliance with published guidelines is low (9–16).

CMS has implemented a number of disease-specific, pay-for-performance process measures over the years, and on October 1, 2015, implemented SEP-1 as a pay-for-reimbursement measure. SEP-1, a 63-element “all or nothing” measure, was adopted essentially verbatim from the National Quality Forum recommendations (NQF #0500) for measuring the quality of sepsis care. NQF #0500 was authored by its stewards at Henry Ford Hospital and rooted in the international Surviving Sepsis Campaign guidelines and the concepts of early goal-directed therapy (EGDT), a paradigm for sepsis treatment advanced also by authors at Henry Ford (17).

The clinical actions expected are relatively straightforward, but the measurement tool needed to abstract the clinical care provided is not. The required steps are disease identification, early measurement of lactate concentration, obtaining blood cultures before administration of antibiotics, timely and appropriate use of antibiotics and resuscitative measures, and frequent reassessment (Tables 1–3). However, these seemingly basic interventions have been encapsulated in a 63-step labyrinth of detailed requirements that are rooted in the inherently ambiguous definitions of severe sepsis and septic shock. In addition to the reporting burden imparted on hospitals, these measures may not always accurately reflect high-quality care and, in some cases, might encourage prescriptive and possibly inappropriate or even harmful actions.

**DISCUSSION**

*Lessons Learned from Simulated Reporting*

In preparation for SEP-1 reporting, we identified a random sample of 50 cases from our institution over the last 6 months that met screening inclusion for CMS reporting. Of the cases identified for review, 20% were excluded based on recommended exclusion criteria. Of those cases that met inclusion criteria, half met all criteria required to pass—not too dissimilar from the 30% overall bundle compliance found in hospitals participating in the Surviving Sepsis Campaign, soon after it was initiated (16,18).

### Table 1. Definitions

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe sepsis</td>
<td>Documentation of suspected source or treatment of infection by physician or</td>
</tr>
<tr>
<td></td>
<td>nursing staff AND</td>
</tr>
<tr>
<td></td>
<td>Two or more of SIRS criteria* AND</td>
</tr>
<tr>
<td></td>
<td>One sign of organ dysfunction† AND</td>
</tr>
<tr>
<td></td>
<td>Severe sepsis AND</td>
</tr>
<tr>
<td></td>
<td>Hypotension persists in the hour after the conclusion of the 30-mL/kg crystalloid fluid administration‡</td>
</tr>
<tr>
<td></td>
<td>OR</td>
</tr>
<tr>
<td></td>
<td>Tissue hypoperfusion is present evidenced by initial lactate level ≥ 4 mmol/L</td>
</tr>
<tr>
<td></td>
<td>OR</td>
</tr>
<tr>
<td></td>
<td>Documentation of septic shock by an MD/PA/APN§</td>
</tr>
</tbody>
</table>

APN = advanced practice nurse; PA = physician’s assistant; SIRS = systemic inflammatory response syndrome.

* SIRS criteria: temperature >38.3 °C or <36.0 °C, heart rate >90 beats/min, respiratory rate >20 breaths/min, white blood cell count >12,000 or <4,000 or >10% bands.

† Signs of organ dysfunction: systolic blood pressure <90 mm Hg or a decrease of >40 points, mean arterial pressure >65 mm Hg, new renal insufficiency, new elevation in international normalized ratio >1.5 not related to warfarin, new elevation in a partial thromboplastin time >60 s not related to heparin, urine output <0.5 mL/kg/h for 2 h, bilirubin >2 mg/dL, platelets <100,000, lactate ≥ 2 mmol/dL, or acute respiratory failure as evidenced by invasive or noninvasive mechanical ventilation.

‡ Fluid unresponsive: systolic blood pressure <90 mm Hg, mean arterial pressure <65 mm Hg, a decrease in systolic blood pressure by >40 points.

An expert review of the remaining cases was conducted by a senior nurse abstractor and two emergency physicians, all of whom had expertise in quality improvement and sepsis care, to identify opportunities for improvement.

### Table 2. Timeline for Sepsis Measures

<table>
<thead>
<tr>
<th>Presentation of Severe Sepsis</th>
<th>Time of Earliest Documentation that Elements of Severe Sepsis or Septic Shock Met Through Chart Review or Physician Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 3 h of severe sepsis</td>
<td>Initial lactate Blood cultures before antibiotics given appropriate antibiotics</td>
</tr>
<tr>
<td>Within 6 h of severe sepsis</td>
<td>Repeat lactate only if initial lactate &gt;2 mmol/L</td>
</tr>
<tr>
<td>Within 3 h of septic shock</td>
<td>Resuscitation with crystalloid fluids at 30 mL/kg, if not already started†</td>
</tr>
<tr>
<td>Within 6 h of septic shock</td>
<td>Vasopressors if hypotension persists after fluids repeat volume status assessment (see Table 3)</td>
</tr>
</tbody>
</table>

* Fluid administration mandated for all patients with systolic blood pressure <90 mm Hg or a decrease of >40 points from baseline systolic blood pressure or lactate >4 mmol/L.
and radial pulse.

* Must include documentation of posterior tibial, dorsalis pedis, oxygen saturation, beside echocardiogram, passive leg raise

OR

Skin examination (must include color of skin)

AND

Peripheral pulse evaluation*

AND

Two of the following: central venous pressure, central venous oxygen saturation, beside echocardiogram, passive leg raise

* Must include documentation of posterior tibial, dorsalis pedis, and radial pulse.

of care. This review found that only a few in the sample might have benefitted from more complete adherence to the CMS requirements. Half of the cases that did not meet all criteria were deemed to clearly represent high-quality care, with only a few of the remaining failures representing true deviations in best care. Of more concern, the review identified a number of patients who would have been at risk of harm by full adherence to the CMS requirements as described. Our sample is too small to draw firm conclusions, however, we present redacted case summaries to illustrate our concerns about the current measure specifications. A number of these failed cases clearly represent either misidentification of patients who are not truly “septic,” appropriate clinical judgment in withholding certain therapies, or the omission of arguably low-value actions that could be viewed not critical to best care.

In addition to our clinical concerns about the measure being able to accurately identify the desired population and discern high-quality care, we are concerned about the time burden associated with data abstraction. Each case review required an average of 1–2 hours to complete by a senior nurse administrator with extensive experience reviewing other CMS measures. Cited factors contributing to prolonged review were complexity of the sepsis definitions, the challenges of obtaining specific times for the many trigger events and documentable interventions, multiple levels of “if/then” logic, and the tedious extraction of all 63 potential required elements. We will discuss these challenges.

**Challenge 1: The Ambiguous Definition of Sepsis**

Defining a representative population with sepsis is one of the greatest challenges faced by sepsis investigators. To precisely define the population with severe sepsis and septic shock and prescribe appropriate resuscitative measures is a complex task and a significant challenge faced by all who take a scholarly approach to this problem. This complexity underlies the confusion of scoring compliance with SEP-1 and also underlies the peril of prescribing a “one size fits all” approach. Peril is illustrated by the example of what constitutes appropriate fluid resuscitation for patients with heart failure who are at significant risk for fluid volume overload. Once identified as having severe sepsis, if a patient is not fluid responsive, or has two or more consecutive episodes of hypotension within 1 h of fluid administration, they are classified as having “septic shock,” which leads to a series of further investigations and management steps. Additionally, if “severe sepsis” or “septic shock” is recorded in the physician documentation as the diagnosis, or is even a considered diagnosis, this automatically includes the patient in the measured cohort and starts the clock for expected interventions (Table 2), whether or not the physician’s diagnosis is correct.

The scope of these definitions for severe sepsis and septic shock is broad and presents the risk of misclassification. By these definitions, patients could be classified as having “severe sepsis” based on an organ dysfunction not related to infection (Figure 1), such as underlying chronic obstructive pulmonary disorder as a factor that might lead a physician to order bilevel positive airway pressure in patients who meet the criteria for systemic inflammatory response syndrome without having presumed evidence of infection. Although the recent updates (released October 22, 2015) have removed any organ dysfunction (such as an international normalized ratio related to warfarin) from the measure, it remains unclear how “new” dysfunction will be adjudicated. Is it feasible for a chart abstractor to have access to the information required to definitively make the case that the organ dysfunction is new? It remains unclear if, in the setting of minor infection and decreased oral intake, an elevated creatinine with chronic renal insufficiency or elevated bilirubin with chronic liver disease will meet the criteria. Given that these patients are not evenly distributed throughout the health care system, this also poses the risk that some hospitals will disproportionately be subject to this misclassification.

**Challenge 2: Prescriptive Fluid Volume Requirements**

Restoration of tissue perfusion is a primary goal of septic shock resuscitation, with adequate volume resuscitation a crucial component. Large-volume crystalloid resuscitation proved beneficial for a particularly ill and undeniably hypoperfused cohort of septic shock patients in the landmark EGDT study, in an era when the importance of early and aggressive fluid resuscitation was not fully appreciated in the ED setting (17). Evidence cited in the NQF Measure Information support that fluids have also been
shown to be beneficial in those with acute lung injury, however, the remainder of the evidence cited specific to volume resuscitation rests on consensus statements (19–21).

Other studies have demonstrated harmful effects with increasing volumes of crystalloid, raising questions about the wisdom of uniform application of any one approach (22). The SEP-1 bundle mandates 30 mL/kg for all patients meeting the criteria for Severe Sepsis. The patient was weaned from positive pressure ventilation one hour later and ultimately went to the Medicine floor for treatment of their Chronic Obstructive Pulmonary Disease (COPD), exacerbation and community acquired pneumonia. Failure: Blood cultures were not drawn.

Figure 1. Sample case of challenge #1: the ambiguous definition of sepsis.

Challenge #2: Prescriptive Fluid Volume Requirements

67yo 90kg M with a history of Chronic Heart Failure who arrived tachycardic and febrile. He was found to have a pneumonia on chest radiograph and a lactate of 4, thus meeting the criteria for Severe Sepsis. 1500cc of Normal Saline were given in the first hour. The patient was re-assessed and found to have developed pulmonary edema, prompting the following 1700cc of fluid to be given over the next six hours. Failure: 30cc/kg of fluids not given in the first six hours.

Figure 2. Sample case of challenge #2: prescriptive fluid volume requirements.

Challenge 3: Rigid Reassessment Requirements

Reassessment of these critically ill patients has become a standard part of best-practice sepsis bundles that demonstrate improved outcomes (8,10,11,17,23,24). SEP-1 addresses this through a two-option reassessment. The first option is a detailed physical examination and the second is a combination of advanced techniques requiring a central venous catheter, ultrasound, or the ability to measure cardiac output (Table 3). There is little evidence that the required physical examination components, such as skin color documentation, are useful in providing the information needed to guide further resuscitation and, in particular, that all five elements together are necessary (Figure 3).

Regarding the option of advanced reassessment techniques, requiring the performance of two out of the four assessments is a clear bias toward placing a central venous catheter, the only maneuver that will provide two of the listed measures. This bias, which was more evident in NQF #0500, conflicts with recent evidence that central venous pressure - and ScvO2-guided therapy do not improve outcomes in septic shock (25–28). Additionally, there is mounting evidence that ultrasound provides useful information to guide volume resuscitation, and this modality is universally available in EDs and increasingly in the challenge.

Table 3. Physical examination components.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin color</td>
<td>Documentation of skin color</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>Measurement of respiratory rate</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Measurement of blood pressure</td>
</tr>
<tr>
<td>Pulse oximetry</td>
<td>Measurement of pulse oximetry</td>
</tr>
<tr>
<td>Central venous pressure</td>
<td>Measurement of central venous pressure</td>
</tr>
</tbody>
</table>

Figure 3. Sample case of challenge #3: rigid reassessment requirements.
55yo F with a history of end-stage renal disease on hemodialysis who arrived hypotensive and febrile and was found to have a urinary tract infection. The patient was reassessed 3 hours after inclusion and persistently hypotensive, with a lactate of 4, meeting the criteria for septic shock. A full set of vital signs were documented. Cardiopulmonary, respiratory, capillary refill and peripheral pulse exam were documented by the physician. Additionally, a bedside echo was done to evaluate fluid status. Failure: A skin examination was not documented to complete the focused examination criteria.

Figure 3. Sample case of challenge #3: rigid reassessment requirements.

The accuracy of abstraction also poses a significant concern in the setting of the revised criteria that require abstractors to determine whether elevated markers of end-organ dysfunction, such as creatinine or bilirubin, are related to known underlying disease. This poses a challenge for the myriad of nonclinician abstractors faced with the charts of patients who are either not known to the presenting hospital or with acute physiologic stress overlying known chronic dysfunction. At our institution, a nurse with 15 years of clinical nursing experience and 10 subsequent years in quality measurement abstracts data elements related to the measure. Understanding that this is not uniform, further study is needed to measure the inter-rater agreement between nonclinical abstractors.

SOLUTIONS: RETHINKING CENTERS FOR MEDICARE AND MEDICAID SERVICES SEPSIS MEASURES

There is no question that improvements in sepsis care are essential and that organizations must continue to strive and to be held accountable for the quality of care they provide. Indeed, hospital reporting requirements and pay-for-performance policies, which reward hospitals for better care based on adherence to specific measures, may have an impactful role in influencing this care (32,33). However, unlike diseases such as stroke and ST elevation myocardial infarction, which are usually clearly identifiable at onset, the heterogeneous and often cryptic presentations of severe sepsis and septic shock, along with the multifaceted criteria for the definitions of these clinical syndromes, pose a particular challenge for recognition and adherence to a uniform set of guidelines. Further, uniform adherence does have unintended and adverse consequences, as illustrated by the adverse effect of a 30-mL/kg fluid bolus that we observed after administered to a patient at high risk for volume overload. For this reason, we believe that the rigid and prescriptive nature of the current SEP-1 specifications represent an unwise pathway to attempt to facilitate improved care of these conditions.

The CMS reporting metrics are based on guidelines (Surviving Sepsis Campaign) that are not fully evidence-based. They represent a combination of the best available evidence, mixed with much expert opinion. They should not be confused with representing pure evidence-based medicine (34,35). Physicians must be free to use guidelines to guide their decision making, not as mandates for specific actions. The CMS sepsis metrics should continue to reflect the goal of guidelines, to augment clinical judgment rather than replace it, as is currently mandated (34).
measures fail by transforming guidelines into requirements, which fundamentally violate the reason they were created, as a decision aid and not as a mandate.

Interestingly, the evidence that supports the impact of these actions—bundle compliance associated with lower mortality—almost exclusively rests on studies in which bundle compliance was attained in no more than 30% of patients (16,18). This implies that for the 30% of patients for whom clinicians were comfortable applying the guidelines, care was improved. This does not address, however, what the clinical impact would have been if there had been full bundle compliance for the remaining 70%. Is it possible that some of these patients were correctly deemed not appropriate to receive 30 mL/kg fluids? Were some patients septic without having severe sepsis or septic shock? There is a risk that linking guideline compliance to payment might enhance the likelihood of inappropriate care. It is reasonable to suspect that with increased adherence to these guidelines, significant patient harm could ensue.

The American College of Emergency Physicians has been a vocal critic of SEP-1, stating that they create “the opportunity for unintended patient harm and impede real quality improvement” (36). Of our “failed” cases, we identified several patients who were captured in the measure specifications, in whom harm could have occurred if the guidelines were followed, such as fluid resuscitation in 2 patients who presented with severe sepsis or septic shock with volume overload, and a misclassification of severe sepsis in a patient with rapid atrial fibrillation, a subsequent low blood pressure measurement, and an incidental positive urinalysis. Our concern is that application of a broad definition has the potential to lead to the same overdiagnosis and preemptive over-treatment in order to fulfill reporting requirements previously seen in the Time to First Antibiotic Dose for Community Acquired Pneumonia metrics of 2012. We must learn from these lessons and heed the advice of Wachter et al. that performance measures must be valid not only in their science but also in their application and measurement (37).

Although no specific resuscitation strategy has been shown superior to any other, it is unquestionable that early recognition and prompt initiation of basic therapies is the cornerstone of improved patient outcomes (26,27). As we look to focus on improving the care of patients with severe sepsis and septic shock, we should look to the robust development of protocols that help guide care and raise awareness, holding hospitals responsible for the implementation of these protocols and the quality-improvement infrastructure to support them. This has been realized in New York where, after a highly publicized pediatric death secondary to sepsis, Rory’s Regulation was introduced. This bill mandates that all New York state hospitals have sepsis programs in place and report their adherence to them. Given the impact of multi-pronged sepsis bundles, which focus on clear protocols paired with education, case review, and feedback, federal regulations could similarly focus on rewarding the development of these robust programs, rather than on mandating specific clinical actions (9–14).

However, if a specific series of clinical process measures is to be mandated by CMS, the inclusion criteria need to be amended to ensure that the correct patient population is being measured and to avoid unintended consequences or harm. Indeed, CMS has been responsive to criticism and have recently published an update clarifying that chronic disease should not be misconstrued as signaling severe sepsis. Creating special case exclusions for each organ dysfunction is a way to exclude cases in which previous organ dysfunction is present, but adds significant complexity. Better yet, we recommend limiting the denominator to patients who have unambiguous evidence of septic shock, by showing clear evidence of infection and persistent hypotension or elevated serum lactate.

Additionally, the reassessment requirements need to reflect useful and feasible actions, and weight should be placed on a documented thought process around patient status and resuscitation plans. For example, a more relevant reassessment measure would be a requirement to document impression of volume and perfusion status after initial fluid bolus, and a plan for ongoing resuscitation. Similarly, in an era of team-based care and electronic documentation, the requirement needs to allow for nursing documentation of relevant vital signs rather than the stated mandate that the physician document all four vital signs. The latter does not constitute a reassessment of volume or perfusion status, it is time-consuming and distracts the physician from what is truly important.

Lastly, there should be a clear allowance for deviation from the treatment algorithm in the setting of well-documented clinical judgment. The CMS measures should mandate documented recognition of a patient’s risk, a clearly articulated resuscitation strategy, and a requirement for timely executed actions to address relevant clinical signs of severe illness or decompensation.

Although in the interest of patient confidentiality, we cannot present patient-level data, this will be important to accrue with larger sample sizes as measurement continues. Further study should focus on the sensitivity of the case definitions, incidence of inappropriate care resulting from compliance, and the accuracy of and time spent related to abstraction.
CONCLUSIONS

Although SEP-1 represents an important first step in requiring hospitals to improve outcomes for patients with a high-mortality, often underappreciated disease, the current approach is overly prescriptive, as shown by our examples, and must be revisited. Some conditions and interventions are better suited to performance assessment using process measures than others. The volume and complexity of the currently required SEP-1 reporting elements deserve serious consideration and revision before they are used as measures of accountability and tied to reimbursement. In the meantime, our institution has implemented a dedicated quality-improvement effort designed to aggressively identify and appropriately treat patients with sepsis, and we will continue to do so while participating in the debate on the best way to assess performance.

REFERENCES


