Infectious Disease Endorsement Maintenance: 0500 Severe Sepsis and Septic Shock Management Bundle

COMMENTS RECEIVED AND MEASURE DEVELOPER RESPONSES

November 28, 2012

NATIONAL QUALITY FORUM

Content

[Comment ID: 2980] Commenter: ACEP QPC, Quality & Performance Committee; Submitted by Mr. Dainsworth Chambers (HPR) 4

Measure Developer Response 5

[Comment ID: 2979] Commenter: Submitted by Mr. Reginald Lavender (SPI 13

Measure Developer Response 13

[Comment ID: 2992; 2991; 2989] Commenter: Cleveland Clinic, Cleveland Clinic; Submitted by Dr. Michael P. Phelan, MD, FACEP (PRO) 14

Measure Developer Response 18

[Comment ID: 2984] Commenter: Submitted by Ms. Carmella Bocchino, MBA, RN (HPL) 27

Measure Developer Response 27

[Comment ID: 2983] Commenter: Kathy Szumanski, ENA; Submitted by Ms. Kathleen Szumanski, MSN, RN, NE-BC (HPR) 28

Measure Developer Response 28

[Comment ID: 2982] Commenter: Submitted by Dr. Christopher Fee, MD (HPR) 29

Measure Developer Response 31

[Comment ID: 2981] Commenter: ACEP QPC, ACEP Quality & Performance Committee; Submitted by Mr. Dainsworth Chambers (HPR) 35

Measure Developer Response 36

[Comment ID: 2978] Commenter: Submitted by Mr. Reginald Lavender (SPI) 40

Measure Developer Response 41

[Comment ID: 2931] Commenter: Submitted by Ron Elkin, MD (Public) 42

Measure Developer Response 43

Summary from the measure developers

More than 2 million patients with sepsis present to emergency departments (ED) each year in the U.S. and carry a mortality rate between 20-49%.[1] In addition, up to 12% of in-hospital cardiac arrests during the first 24 hours of admission carry an admission diagnosis of pneumonia which in actuality was sepsis.[2] The average duration of an ED stay in the U.S. is 6 hours for a septic patient.[3] Sepsis is the most common admission diagnosis in U.S. hospitals and patients are 8 times more likely to die of sepsis in the hospital than any other admission diagnosis. More than $62 billion per year of Medicare and Medicaid funds are spent on this disease making this a high priority for CMS and other third party payers.[4]

The good news is that during the last decade we have shown that taking advantage of the first 6 hours can make a significant impact. One of every 5-6 patients can be saved, a decrease in 5 hospital days per admission and a 20% reduction in sepsis related costs can be realized by applying the proposed sepsis composite measure.[5] Thus, improving outcomes and decreasing health care resource consumption from sepsis should be an immediate concern for all specialties involved in the care of these patients. Rather than penalize individuals, the proposed composite measure promotes cooperation between clinicians and hospitals to improve care.

While the ED is the portal of hospital entry for 52% of patients with severe sepsis, the remaining 48% come from the general practice floors and the ICU. Thus, this proposed sepsis measure is not a specialty-centered measure but rather a hospital-wide measure. To provide the best outcomes, the measure should be applied along with a robust continuous quality improvement initiative that engages parties in all locations throughout the hospital. The composite measure is the key to success in such an initiative.

In the evolution of improved outcomes due to AMI, stroke, and trauma a coordinated approach between the ED and hospital specialties was necessary to improve outcomes. Hospitals and professional specialties responded as a result of the potential benefits for patients. We appreciate the opportunity provided by the National Quality Forum (NQF) to continue this type of work as regards sepsis care with our colleagues from various national professional societies who we know have the best interests of patients in mind.

Submitted by ACEP QPC, Quality & Performance Committee; Submitted by Mr. Dainsworth Chambers

The American College of Emergency Physicians (ACEP) Quality and Performance Committee would like to thank the Steering Committee for the opportunity to comment on the proposed changes to the NQF Measure #0500: Severe Sepsis and Septic Shock Management Bundle. ACEP believes that all of our septic patients deserve timely treatment, which is the hallmark of emergency care. In fact, mortality is significantly reduced for septic patients who present to the ED compared to those who are admitted directly to the ICU or the floors, because of the timely high quality care treatment they receive (Powell 2012). However, with many EDs caring for patients being boarded for hours and days without an inpatient bed, we question the face validity of using triage time as time zero. Many ED patients will present with uncomplicated pneumonia, urinary tract infection, or cellulitis only to meet the criteria for severe sepsis/septic shock hours later. If the measure calls for early goal directed therapy within three hours of triage, but the patient does not meet criteria for severe sepsis or septic shock until four hours later, then even if all required interventions are completed within an hour, the hospital will fail on this measure as currently specified. That type of measurement does not differentiate hospitals based on the quality of care provided, but rather on the ED length of stay. If used for accountability as specified, this measure could cause the unintended consequence of penalizing large volume and safety net hospitals.

The concept of a timed accountability measure, which uses a symptom-based assessment at triage as time zero, yet does not exclude all patients who develop symptoms subsequent to triage makes reliable measurement impossible. The measure developer response that patients presenting to the ED with signs of infection are “somewhere on the natural trajectory of becoming septic regardless of point of presentation” is unacceptable for an accountability measure. Furthermore the claim that this measure could be retrospectively reviewed based on ICD-9 hospital discharge codes, with triage as time zero is not an accurate measure of quality of care provided in the ED. Previous accountability measures attempting to measure the time to antibiotics for patients with pneumonia have already shown poor fidelity when using ICD-9 discharge codes. Furthermore, application of the NQF’s “Guidance for Measure Testing and Evaluating Scientific Acceptability” would

classify the reliability of this measure as “low” when “one or more measure specifications are ambiguous with potential confusion in identifying who is included and excluded from the target population, or the event, condition, or outcome being measured; or how to compute the score.” None of the reliability data submitted addresses the core question of if the triage time is a reliable measure of when severe sepsis or septic shock starts.

Reply to Chambers from the measure developers:

We would like to thank Mr. Dainsworth Chambers of the ACEP Quality and Performance Committee for his comments. These inquiries and comments closely resemble those provided by additional members of ACEP and Society of Academic Emergency Medicine (SAEM) later is this document. In comprehensively addressing Mr. Chambers’ inquires and comments we hope to provide answers to similar questions posed by his colleagues.

1. Mr. Chambers states, “…with many EDs caring for patients being boarded for hours and days without an inpatient bed, we question the face validity of using triage time as time zero.” This appears to be a comment reflecting a threat to validity, specifically “face validity.”
    a. With all due and proper respect, it is necessary to note that Mr. Chambers offers no data as evidence for a threat to validity. This absence of evidence stands in contrast to the measure developer’s provision of statistically significant data establishing, in the maintenance committee’s opinion, a moderate degree of validity.
    b. By NQF standards, “face validity” is the least acceptable evidence of validity. While Mr. Chambers raises a concern about face validity, the question does not amount to a demonstration that the use of triage time is invalid. The measure developer has provided evidence of validity exceeding face validity alone in the submission.
    c. It is not clear to the measure developers how a concern for patients “being boarded for hours and days without an inpatient bed” in an ED prevents the identification and treatment of severe sepsis in the ED. Boarding of this sort is certainly not a threat to face validity since this long period of time seemingly would permit identification and treatment in the ED, rather than preventing good care.
    d. There is precedent for triage time in acute illness. It is currently used as time zero for trauma patients. Whether major or minor, this time point is the national standard for evaluation and treatment

of trauma in the United States. Illness severity frequently changes during the ED stay for this disease.

e. Risk stratification in early sepsis detection is one of the key aspects of a quality improvement strategy with measurement. The presence of hypotension[6] or lactate greater than 4 mM/L[7-11] alone is associated with a mortality in excess of 30%. When both of these findings are present at triage or develop during the ED stay, mortality exceeds 46%. This rate of mortality was seen in the original EGDT study[12] and has been confirmed by the Surviving Sepsis Campaign (SSC) in greater than 15,000 patients a decade later.[13] It only makes sense that early detection at triage or the most proximal stage of hospital presentation of these high risk patients provides the most optimal benefit.

f. Defining triage time as time zero is an objective starting point that removes the variability of individual clinicians' abilities to detect severe sepsis. Setting triage time as time zero allows for continuous process improvement to improve early identification. Since many hospitals currently lack a screening program, this time point has proven to be most effective at moving hospitals toward early detection. This was comprehensively and rigorously shown to be valid and reliable in the SSC database in more than 28,000 patients as presented in the submission.[13]

2. Next Mr. Chambers states, "[i]f the measure calls for early goal directed therapy (EGDT) within three hours of triage, but the patient does not meet criteria for severe sepsis or septic shock until four hours later, then even if all required interventions are completed within an hour, the hospital will fail on this measure as currently specified." Again, with all due respect, this position is inaccurate for several reasons:

   a. The measure calls for EGDT by 6 hours of triage, not 3 as stated above. Thus, the patient Mr. Chambers describes would have 2 additional hours to meet measure requirements.

   b. Failure to meet all the components of the measure in a patient who qualifies whether in the ED, general wards or ICU is the overall objective standard. For the ED specifically, the measure developers have provided data that in more than 200 hospitals when triage time has been used as time zero, a steady increase in compliance and reduction in mortality has been universally achieved.

c. Mr. Chamber's concerns should perhaps be tempered by the fact that outcomes are improved when the measure is applied.

3. Mr. Chambers notes, "If used for accountability as specified, this measure could cause the unintended consequence of penalizing large volume and safety net hospitals."
    a. The concern raised here regarding "accountability" apparently refers to the adoption of the measure by third-party payers with financial penalties attached to the measure, such as CMS' use of value-based purchasing measures in recent years. Mr. Chambers (and others who raise this objection throughout the public comments) misapprehend how these measures have been applied by such payers. All such measures employed by CMS to date have defined percentiles of compliance. Thus, if the best any hospital can do is achieve a 35% compliance rate with the measure, that hospital falls in the P90 or above for the metric. CMS applies penalties for performance below P50 presently in value-based purchasing metrics. There is no reason to believe that Sepsis #0500 would be treated any differently as an "accountability" metric. Thus, since compliance in such a scheme will be calculated relative to all hospitals' performances (as a percentile) there is little to fear from an individual hospital's difficulty in complying with the metric. In fact, the metric would perform as intended in such a circumstance, driving efforts to increase compliance, which the provided evidence supports.
    b. The evidence provided from more than 200 hospitals in the Surviving Sepsis Campaign database does not fail to include "large volume hospitals." For example, the data includes large university hospitals such as Rhode Island Hospital in Providence, Rhode Island as well as Cooper University Hospital in Camden, New Jersey. The trend toward increasing compliance over time was proven in these populations. No data are provided by Mr. Chambers to support his assertion that large volume hospitals may be penalized.
    c. "Safety net hospitals" have been excluded from penalties associated with value-based purchasing metrics. Moreover, there is no evidence that safety net hospitals cannot improve performance under the proposed measurement system.

4. Mr. Chambers states, “The concept of a timed accountability measure, which uses a symptom-based assessment at triage as time zero, yet does not exclude all patients who develop symptoms subsequent to triage makes reliable measurement impossible” (emphasis supplied). Mr. Chambers frames this comment as a threat to the reliability of the measure.
    a. The comment reflects misunderstanding as to the nature of reliability in measurement. NQF has stated that, “Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise.” Here, the measure developer has provided evidence that when triage time is used as time zero for patients presenting to the ED, more than 200 hospitals are able to produce the same results a high proportion of the time. Indeed, the committee ranked the reliability of the measure as having the highest score possible based on the NQF approved (see Measure Testing Task Force Final Report) RAND methodology. Use of the RAND methodology permitted computation of performance scores for the hospitals in the SSC database and a demonstration of reliability using a signal-to-noise analysis. The results demonstrated unequivocally that the metric functions with the highest degree of reliability.
    b. We must respectfully state the premise of Mr. Chambers’ claim is incorrect: “…a symptom based assessment at triage as time zero” is not a requirement of the measure. Rather, the act of triage itself, whether symptoms are assessed or not, is the definition of time zero in the ED. Symptoms do not factor into setting the triage time. In fact, because the triage time is a hard data point, reliability as a property of a measure is greatly enhanced. This means the exact opposite of Mr. Chambers’ intimation is true as regards the reliability of this measure – defining time zero as detection of the signs and symptoms of severe sepsis would be less reproducible, less reliable than a hard data point such as triage time. It would introduce the variability of time of detection (or worse clinicians’ documentation) which is highly clinician dependent.

c. We submit that Mr. Chambers has the burden of demonstrating that the measure is unreliable when strong evidence suggests the opposite is true. The absence of evidence to support this claim renders Mr. Chambers' larger position problematic.

5. Mr. Chambers takes issue with, "[t]he measure developer[s'] response that patients presenting to the ED with signs of infection are 'somewhere on the natural trajectory of becoming septic regardless of point of presentation'" as "unacceptable for an accountability measure."
   a. These comments are again made in the context of Mr. Chambers' concern with the reliability of the measure. We wish to reiterate that the measure is reliable as evidenced by statistical analysis in more than 200 hospitals and as noted in 4a above.
   b. Mr. Chambers' comment seems to reflect concern that there is variation between the time of triage in the ED and the time the patient is *in actuality* septic. Variation itself is not necessarily a deadly threat to the reliability or validity of a measure. On the other hand, variability may be great enough to render a measure unreliable or invalid. This situation is resolved in the science of measurement through the use of statistical tests. Here, the calculations provided in the Sepsis #0500 submission account for the variation that concerns Mr. Chambers. The results conclude quite convincingly that the degree of variation is not a threat to reliability or validity. When a measure developer has accounted for the variation with a proper statistical demonstration of reliability and validity the measure should be approved. It is incumbent upon Mr. Chambers or others to demonstrate the variation is indeed a threat, which it does not appear to be in this instance. We respectfully submit Mr. Chambers has brought no evidence in this regard.
   c. Given that the degree of variation is within proper statistical limits, the measure is appropriate as an accountability measure. The measure developer does not assert that more precise measures are impossible in the future. The evidence now available, suggests the measure is sufficiently precise as to be reliable and valid.
6. Mr. Chambers notes, "…the claim that this measure could be retrospectively reviewed based on ICD-9 hospital discharge codes, with

triage as time zero is not an accurate measure of quality of care provided in the ED."

a. The statement appears to suggest that ICD-9 codes might somehow fail to identify patients with severe sepsis and septic shock. There is no reason to believe this is true. Use of ICD-9 codes at discharge is presently the standard by which CMS identifies many patient populations such as heart failure patients or acute myocardial infarction (AMI) patients. As regards measures associated with AMI, for instance, the list of patients with a discharge diagnosis for AMI is generated quarterly from ICD-9 codes and a sample of those patients is selected for chart abstraction. Those charts are then reviewed by quality department abstractors and compared to the CMS measure specifications for items such as "door to balloon time" if the patient presented to the ED with AMI, for instance. In the case of severe sepsis, the same routine would apply. These abstractors would determine if severe sepsis was present in the ED and if so, time zero would be defined as triage time. There is nothing inconsistent with this strategy for severe sepsis and present strategies for other diagnoses.
b. Mr. Chambers comments that the use of ICD-9 codes would not be an "accurate measure of quality of care provided in the ED," but he cites no reason as to why the use of these codes would fail to reflect the care provided in the ED. Moreover, the measure reflects hospital care in general, applying to patients presenting from other venues as well. In those instances where severe sepsis is not present in the ED, ED triage time would not be the standard applied, ensuring that misattribution is not made to the ED.
c. If Mr. Chambers' statement was meant to reflect that retrospective review based on ICD-9 codes is not possible, we respectfully suggest that the experience of hospitals doing so in the Surviving Sepsis Campaign database would counter that position.
d. It should be noted that in the maturation of a sepsis continuous quality initiative, a screening strategy is almost always adopted. Thus, many more patients will be detected concurrently with less reliance on retrospective diagnosis. Although measurement may be still be based on retrospective data, a mature quality improvement process will know who these patients are in advance.

7. Finally, Mr. Chambers states, "…application of the NQF's 'Guidance for Measure Testing and Evaluating Scientific Acceptability' would classify the reliability of this measure as 'low' when 'one or more measure specifications are ambiguous with potential confusion in identifying who is included and excluded from the target population, or the event, condition, or outcome being measured; or how to compute the score.' None of the reliability data submitted addresses the core question of if the triage time is a reliable measure of when severe sepsis or septic shock starts."
    a. We must respectfully point out that there is no 'confusion in identifying who is included and excluded from the target population.' The measure specifications specifically define inclusion criteria as those patients who meet criteria for severe sepsis and septic shock. The specific criteria require documentation of a suspected source of infection, the presence of 2 or more variables indicative of systemic inflammatory response syndrome (SIRS), and meeting stated variables consistent with organ failure. Therefore, if a chart abstractor pulled a chart based on an ICD-9 code for instance, and those criteria were not met, the patient would be excluded.
    b. We respectfully believe that Mr. Chambers conflates his dissatisfaction with use of triage time as time zero with the entirely distinct specifications for inclusion and exclusion in the measure. To be clear, one issue is about starting the clock for the measure, the other is about who is included in the measure. We are forced to point out that there is no linkage between them. To expressly make the point, the triage time would be irrelevant if the chart failed to meet inclusion criteria. Thus, the rating of high reliability remains appropriate. The confusion Mr. Chambers references simply is not truly present.
    c. Regarding the specific concern that "[n]one of the reliability data submitted addresses the core question of if the triage time is a reliable measure of when severe sepsis or septic shock starts," Mr. Chambers has mistakenly concluded that the measure developer asserts that triage time and the start of severe sepsis or septic shock are the same. We do not consider triage time a proxy for the start of severe sepsis or shock. We do consider triage time a reliable data point to initiate a measurement for patients who

happen to develop severe sepsis or shock during their stay in the ED. We must respectfully state that the belief that we have rendered the times as *equivalent* is plainly in error.

d. Moreover, we assert (with statistical support) that any variability present between ED triage time and the actual start of severe sepsis and septic shock in the ED is not so great as to threaten the reliability and validity of the measure. In fact, quite the opposite is true. The association is sufficiently tight that the degree of reliability is high. Our evidence and statistical reasoning show convincingly that if measurement begins at triage for patients presenting with severe sepsis or septic shock during their ED stay, hospitals are able to comply with the measure and improve their compliance over time while decreasing mortality. Mere assertion that the two time points are not linked is insufficient evidence compared with the data presented by the measure developer.

Submitted by Mr. Reginald Lavender

Edwards appreciates the opportunity to provide comments on the Infectious Disease Consensus Standards Endorsement Maintenance 2012 Draft Report. Edwards has been a world leader in advanced cardiovascular treatments for the last forty years. Edwards is the global leader in acute hemodynamic monitoring and the number-one heart valve company in the world. Headquartered in Irvine, California, Edwards offers medical technologies for debilitating and life-threatening conditions, including brands with leading global market positions such as SAPIEN, FloTrac, Fogarty, PERIMOUNT, and Swan-Ganz. Our products care for patients of all ages undergoing surgical cardiology procedures, especially in complex cases with severe co-morbidities requiring specialized care.

Edwards commends NQF for prioritizing infectious diseases, specifically sepsis and septic shock, measures. Edwards supports the efforts to improve the quality of care by means of measure development and endorsement to sustain the movement towards implementing a high-value healthcare system. As NQF considers the direction of future measure endorsement, increased attention and focus on sepsis and septic shock should be prioritized. Edwards appreciates NQF's request for more information and data to further consider #0500: Severe Sepsis and Septic Shock: Management Bundle for endorsement.

Reply to Lavender from the measure developers:

We appreciate the commentary and support for this measure.

Submitted by Cleveland Clinic, Cleveland Clinic; Submitted by Dr. Michael P. Phelan, MD, FACEP

Although we support the quality improvement efforts of the Surviving Sepsis Campaign, we have concerns regarding the scientific acceptability, validity/reliability and feasibility of data collection of NQF modified Measure #0500: Severe sepsis and septic shock management bundle for use in any public reporting measure. The measure will likely need to be an abstracted measure and the current definition will be an extremely difficult and time consuming process unless the measure is significantly modified. Specifically we have concerns about:

1) The reliability of triage being time zero for ED patients

2) Lack of evidence for the central venous pressure (CVP) measure component or blood culture prior to antibiotics elements

3) Feasibility and data abstraction burden on hospitals of an 8 component composite measure with timed elements.

Concerns about some of the metrics specifically:

The definition of time of presentation. For ED patients utilizing time of "initial triage " rather than "earliest chart annotation consistent with all elements of severe sepsis…" is unfair and a single unified definition should be utilized (i.e. "earliest chart annotation consistent with all elements of severe sepsis…"). Often time's patient present to the ED with normal vital signs then decompensate and meet criteria of sepsis. Including the initial time of presentation as the start time may not reflect patient's condition adequately. This ambiguity of utilizing different criteria of time of presentation based on location, calls into question the measure reliability. NQF's Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties would classify the reliability of this measure as "Low" when "one or more measure specifications are ambiguous with potential confusion in identifying who is included and excluded from the target population, or the event, condition, or outcome being measured; or how to compute the score." Therefore, this measure does not in fact meet the NQF criteria for reliability. Simplifying the time of presentation to "earliest chart annotation consistent with all elements of severe sepsis ascertained through chart review" and

eliminating the “time of triage in the emergency department (ED)” would improve the metric.

At least 2 of the 7 composite measures in the numerator statement should be either eliminated due to it not being a validated element and potential for unintended consequences or modified.

F. In the event of persistent arterial hypotension despite volume resuscitation (septic shock) or initial lactate >=4 mmol/L (36 mg/dl) measure CVP and central venous oxygen saturation

B. Obtain blood cultures prior to antibiotics

There may be the unintended consequence of increasing the use of central lines in situation where they may actually not be needed and potentially causing harm by their placement (bleeding pneumothorax, pain) or causing infections. By including this single item in the composite measure may encourage the over utilization of central line placement specifically not to fail the measure rather than taking care of the patients best interests. This item forces one to either measure CVP or $ScvO_2$ by mandating a central venous line be placed. The scientific acceptability of the practice needs to be better explored since neither of these practices has been reliably shown to effect outcome.

Moreover, NQF’s own Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety – Composite Measures clearly states in Table 1 that “the individual measures included in the composite or sub composite must be either NQF endorsed; or assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure.” Currently there are no NQF-endorsed measures that address CVP in septic patients. Several of the new component indicators including CVP where not actually included in the current NQF endorsed measure #0500, because they did not meet the NQF criteria for scientific acceptability as component measures at that time they were endorsed.

The multiple other ways to assess volume status other than invasive CVP measurement of SVO_2, especially non invasive methods like bedside echocardiography or IVC measurement have not been adequately explored or added as an exception. However we would strongly recommend that

element F just be eliminated from the language of this metric until better data shows this action item to be associated with actual improved outcomes. Or if needed it can be added as a non contributory element of the measured composite (rather than contributing to the actual composite score).

We also would consider eliminating element B or modifying the wording; “obtaining Blood cultures prior to antibiotics”. The whole point is that the patients receive broad spectrum antibiotics not that they are timed prior to antibiotic administration. The theoretical concern about sensitivities should not trump actual administration of those antibiotics. If not eliminated than perhaps altering the wording to simply state; “obtaining appropriate cultures” which would allow simplicity and more flexibility in the actual abstraction process. Having to identify the time of antibiotic administration along with the time of collection of cultures adds significantly to the burden and complexity of the abstraction process. Theoretically this may seem important but does the act of obtain blood cultures or any culture prior to the administration of antibiotics actually have any effect on outcomes?

Feasibility and data abstraction burden on hospitals of an 8 component composite measure with timed elements.

While the currently NQF-endorsed measure only addresses the initial four components (lactate, blood cultures, antibiotics, fluids) for severe sepsis, the new proposal includes an entire composite of its own with three additional components (vasopressors, CVP, and ScvO2) for septic shock. This new composite is far too complex for implementation as a potential accountability measure. Furthermore, all of the data elements and time stamps required to calculate this measure are not readily available discrete fields from existing electronic sources making it a significant burden on hospitals to sort and collect this data. Perhaps evens simple baseline composite tiering system to score this composite metric would work with the 3-4 most important elements of the bundle included which are already included in the original NQF endorsed sepsis metric:

1. IVF’s administered (yes/no)

2. antibiotics administered (yes/no)

3. lactate measurement (yes/no)

4. cultures (yes/no)

Because in all likelihood this metric will be an abstracted measure, the simplification of the language and not hardwiring failure maybe be the best way to get a sepsis measure up and running that will be understandable and useable. Perhaps if 3 of the 7 elements are met that could be the cut off for passing the metric instead of an all or nothing approach.

Other elements that would not contribute to the composite metric could still be collected but utilized for research /data mining purposes i.e. to evaluate their efficacy or validity.

Prior to accepting a measure like this we would like to see data abstraction algorithm and testing of the data collection process to assess its data collection burden and feasibility.

Reply to Dr. Phelan from the measure developers:

We appreciate the opportunity to address the thoughtful comments of Dr. Phelan.

Dr. Phelan has identified 3 major concerns. These concerns are similar to those of Mr. Chambers and we will reference some of our answers to both throughout.

At the outset, it is necessary to point out that Dr. Phelan has included a number of suggested alterations in the measure. While we appreciate his initiative in this regard, the obvious problem is that these suggestions for the most part cannot be adopted because no testing has been done to establish the effectiveness of the altered measures he suggests. The NQF framework requires specific testing of measure elements. The evidence base for these spontaneous suggestions is simply lacking.

Dr. Phelan's first concern is the definition of the time of presentation. He makes several points that will be individually addressed.

1. Dr. Phelan notes that, "For ED patients utilizing time of 'initial triage' rather than 'earliest chart annotation consistent with all elements of severe sepsis…' is unfair and a single unified definition should be utilized (i.e. 'earliest chart annotation consistent with all elements of severe sepsis…)." Dr. Phelan's position here is that the definition is in some way "unfair."
    a. We must respectfully state it is not clear in what sense the measure is "unfair." Indeed, the evidence provided in the submission for the use of triage time in the ED along with the use of earliest chart annotation consistent with all elements of sepsis on the wards remains unchallenged. The evidence we present in the submission has met the standard of peer review in multiple publications, statistical significance in terms of correlation with mortality reduction, and finally the reliability and validity standards set by NQF. We respectfully suggest Dr. Phelan has provided no evidence of unfairness in the strategy that would be needed to refute the high bar the measure developer has already met. Simply asserting that the measure is in some way unfair does not make it so. To understand the

objection one would need to know to whom the measure is unfair, in what way, if the alleged unfairness is statistically demonstrable, etc.

2. Dr. Phelan next notes that, "Often times patient[s] present to the ED with normal vital signs then decompensate and meet criteria of sepsis. Including the initial time of presentation as the start time may not reflect patient's condition adequately."
    a. We must respectfully note that Dr. Phelan here invokes the same incorrect reasoning cited by Mr. Chambers that we, the measure developers, believe that the triage time is identical to the point in time that the patient is *in actuality* severely septic. Please see our refutation of this position in section 7b and 7c above. In short, we do not assert that the times are the same. We assert only that for purposes of measuring the care of patients who become severely septic in the ED, a reliable time to start the clock to facilitate measurement is triage time.
    b. Nevertheless, we respectfully believe it is beyond coincidence that patients who present to the ED and develop documentable signs and symptoms of severe sepsis during their stay presented to the ED for anything other than their imminent decompensation due to severe sepsis. Quite plainly, it is not as if such patients present with common unrelated diagnoses such as a fractured limb and then, *par coincidence*, slip into severe sepsis. We believe that the assertion that there is no correlation between the presentation and the eventual identification of severe sepsis should offend the committee's reasonable sensibilities. The far more common issue is that poor systems exist to identify severe sepsis in the ED when it should have been clinically apparent or highly suspected. Such deficiencies are highly amenable to quality improvement projects. The less frequent scenario Dr. Phelan envisions should not govern the far more common systems level issues, especially where evidence has been submitted demonstrating improvement is made possible with this composite measure.
    c. In any event, any suggestion of excessive variation imposed by the use of triage time that would threaten reliability or validity of the composite measure is unfounded. We have established that

even with the degree of variation Dr. Phelan has identified, the measure remains statistically sound meeting NQF criteria for both reliability and validity. Please see our answer in 7C above and the statistical detail in the original submission.

3. Dr. Phelan raises a concern that, "This ambiguity of utilizing different criteria of time of presentation based on location, calls into question the measure reliability. NQF's Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties would classify the reliability of this measure as "Low" when "one or more measure specifications are ambiguous with potential confusion in identifying who is included and excluded from the target population, or the event, condition, or outcome being measured; or how to compute the score." Therefore, this measure does not in fact meet the NQF criteria for reliability."
    a. Dr. Phelan thus suggests that differing standards for the ED (time of triage) and the wards (earliest chart annotation consistent with all elements of severe sepsis) make the measure ambiguous. However, we respectfully submit that it does not logically follow that just because the standards are different they are 'ambiguous' or result in 'potential confusion in identifying who is included and excluded from the target population.'
    b. We must point out for clarity that there is, in fact, no actual ambiguity: if the patient develops severe sepsis or septic shock during the stay in the ED, triage time is used. If the patient develops severe sepsis or septic shock on the hospital wards, the earliest chart annotation consistent with all elements of the diagnosis is used. Any abstractor would be capable of drawing the distinction by chart review. If perhaps Dr. Phelan's concern is that abstractors may be incompetent to draw the necessary distinction, CMS (for instance) validates the competencies of hospitals' chart abstractors yearly through the local Quality Improvement Organization (QIO).
    c. There is precedent for this type of distinction in chart abstraction for national quality measures. For instance, patients presenting to the ED with ST elevation MI are measured for "door to balloon time." Patients who develop ST elevation on the wards are not abstracted for the ED based measure. Rather, a patient on the

wards will be measured from time of EKG demonstrating ST elevation to percutaneous intervention or thrombolysis. Quality professionals who engage in chart abstraction make determinations about which measure to apply based on location routinely.

d. Therefore, we respectfully note that Dr. Phelan's concern that the measure 'does not meet NQF criteria for reliability' due to ambiguity or some sort of confusion is unsupported and in fact at odds with precedent.
e. Dr. Phelan proposes that "Simplifying the time of presentation to 'earliest chart annotation consistent with all elements of severe sepsis ascertained through chart review' and eliminating the 'time of triage in the ED' would improve the metric." We respectfully must also point out that this is not the case. The reliability that is garnered by a hard data point such as triage time contributes to the robustness of the metric in terms of the reliability calculations submitted in Sepsis #0500. Removing this time would perhaps be more likely to render the metric unreliable.

4. Dr. Phelan next suggests that element F in the numerator statement (*In the event of persistent arterial hypotension despite volume resuscitation (septic shock) or initial lactate >=4 mmol/L (36 mg/dl) measure CVP and central venous oxygen saturation*) should be removed.
    a. In favor of this suggestion, Dr. Phelan remarks that the use of central lines in situations "where they may not actually be needed" will cause potential harm, and that "This item forces one to either measure CVP or $ScvO_2$ by mandating a central venous line be placed. The scientific acceptability of the practice needs to be better explored since neither of these practices has been reliably shown to effect outcome."
        i. We respectfully submit that the statement that the line "may not actually be needed" does not acknowledge the strong evidence that mortality is reduced when patients receive EGDT. The extensive evidence in favor of EGDT has already been adjudicated by NQF passed the test of

importance and scientific validity in the maintenance committee.

ii. In regards to harm, we note that Dr. Phelan has not quantified the likelihood of such harm but instead has remarked on it as a theoretical possibility. Such harm is the exception rather than the rule in line placement. To defeat the power of the evidence provided in the submission, a great degree of evidence would be required to demonstrate that the harm outweighs the benefit.

iii. There is strong evidence early CVP placement is associated with improved organ dysfunction. Kashiouris et al, identified 341 patients who had received a CVC within 24 hours after sepsis recognition for EGDT. Every additional hour of delay from sepsis recognition to the first CVP measurement was associated with an average of 0.14 additional point increase in the 2-day delta SOFA score on (95% CI 0.10–0.18 p <0.002). Thus, in patients with severe sepsis and septic shock, every additional hour of delay in CVC utilization was associated with worsened multi-organ failure in the first 48 hours after ICU admission.[14]

iv. There is strong evidence early CVP placement is associated with improved organ dysfunction.[15] Walkey et al analyzed the proportion of septic shock cases receiving an early (day of admission) CVC and the odds of hospital mortality. From 1998-2009, 203,481 (population estimate: 999,545) cases admitted through an ED with principal diagnosis of septicemia and secondary diagnosis of shock were examined. From 1998-2009 population-adjusted rates of septic shock increased from 12.6 cases per 100,000 US adults to 78 cases per 100,000. During this time age-adjusted hospital mortality associated with septic shock declined from 40.4% to 31.4%. Early CVC insertion increased from 5.7% (95% CI 5.1-6.3%) to 19.2% (95% CI 18.7-19.5%) cases with septic shock, with an increased rate of CVC utilization identified after 2007. The rate of decline in age-adjusted hospital mortality was significantly

greater for patients who received an early CVC (-4.2% per year, 95% CI -3.2, -4.2%) as compared with no CVC (-2.9% per year, 95% CI -2.3, -3.5%), p=0.016. Hospital mortality associated with early CVC insertion significantly decreased from a multivariable-adjusted odds ratio of 1.29 (95% CI 1.14-1.45) prior to 2001 to an adjusted odds ratio of 0.87 (95% CI 0.84-0.90) after 2001. Utilization of a CVC early in septic shock has increased 3-fold since 1998.[16]

v. These above studies reveal that organ failure and mortality associated with early CVC insertion has decreased after the EGDT publication in 2001.

b. Dr. Phelan next requests consideration that element F should be removed because NQF's Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety–Composite Measures states that "the individual measures included in the composite or sub composite must be either NQF endorsed; or assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure." Dr. Phelan notes that there are no component measures endorsed for CVP at this time.

i. We are respectfully compelled to point out that Dr. Phelan has left out the very next sentence in the report. It goes on to state, "A component measure might not be important enough in its own right as an individual measure, but it could be determined to be an important component of a composite." Such is the case in this instance. The composite indicator has been extensively evaluated with the inclusion of CVP. In each case cited in the submission, as well as in the analyses done for reliability and validity and feasibility, CVP was included in the overall measurement strategy.

ii. Indeed, it is the case that we do not know how the composite measure would perform absent the component of CVP since in each analysis the metric was included. We believe, therefore, following Dr. Phelan's recommendation

would thwart the entire strategy rather than allowing for simple deletion of item F as he requests.

iii. We the measure developers suggest, following the NQF reasoning, CVP as an individual measure would not be important enough to separately endorse, but it plays a key role in the composite permitting some assessment of intravascular volume from which the clinician may draw inferences about further care. None of the methods of assessing intravascular volume that Dr. Phelan has referenced has proven to be more efficacious than another across patient presentations. We are compelled to point out that Dr. Phelan provides no evidence that other indicators of intravascular volume are universally regarded as superior strategies.

iv. Finally, the committee evaluating the evidence during the maintenance cycle fully considered all of the available evidence in support of and against the proposed changes. The committee concluded that the proposed changes meet the required scientific evidentiary standards. The submission included sub analyses of the component measures with respect to reliability and validity as well. In each instance, the committee approved the composite. The concern raised here is not an adequate substitute for the committee's judgment.

c. Dr. Phelan's suggests another reason to consider removal of element F. He notes there are "multiple other ways to assess volume status other than invasive CVP measurement of $ScvO_2$".

i. The measure developer wishes to clarify that $ScvO_2$ is a measure of tissue oxygenation rather than a measure of volume status as implied.

ii. Please note that the above statement takes no issue with the use of $ScvO_2$ as a measure of tissue oxygenation and, as it happens, central venous access is required to measure $ScvO_2$. Thus, obtaining CVP from that catheter is no added burden and carries no further risk of harm.

iii. To the extent that Dr. Phelan is concerned that other means of assessing intravascular volume will not be

utilized, nothing in the measure prohibits clinicians from assessing intravascular volume by other means.

5. Dr. Phelan next requests consideration that element B in the numerator statement (*Obtain blood cultures prior to antibiotics*) should be removed or that the wording should be modified.
    a. Dr. Phelan's reasoning is that "The theoretical concern about sensitivities should not trump actual administration of those antibiotics." The Surviving Sepsis Campaign agrees with the spirit of this conclusion in the 2012 guidelines. The recommendation is not to delay antibiotic administration beyond 45 minutes waiting for cultures to be obtained. A specification in that regard could easily be added to the final measure.
    b. Dr. Phelan cites a concern with the difficulty of abstraction by looking for time stamps in the clinical documentation. We respectfully submit that quality professionals by now are familiar with the requirement to locate time stamps for culture collection since these elements are included in CMS measures such as pneumonia and the Surgical Care Improvement Project (SCIP).
    c. Dr. Phelan asks the question, "does the act of obtain[ing] blood cultures or any culture prior to the administration of antibiotics actually have any effect on outcomes?" The answer appears to be "yes." Choosing the wrong antibiotic can lead to a five-fold increase in mortality, thus this step is potentially fatal to patients.[17] Correcting an error with culture data may be life saving.
6. Dr. Phelan next addresses "the feasibility and data abstraction burden of an 8 component measure with timed elements." He remarks that, "This new composite is far too complex for implementation as a potential accountability measure," and that "all of the data elements and time stamps required to calculate this measure are not readily available discrete fields from existing electronic sources."
    a. We respectfully submit that we are quite confident that this data abstraction burden is not uniquely burdensome to hospitals. For instance, the Surviving Sepsis Campaign database's data elements which yielded more than 28,000 patients' data was feasibly collected at more than 200 hospital. In another example, CMS has already exceeded the burden of an 8

component measure with the SCIP measure set. That measure set requires hand abstraction of 32 individual discrete measures. Please see the latest version of the SCIP specifications available from Quality Net at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228772433589.

b. As regards the concern that not all data elements are readily available from discrete electronic sources, this may or may not be true, but is in any event not unique to Sepsis #0500. The vast majority of the 5000 hospitals across the United States do not have an electronic health record. The standard for abstraction across the United States for CMS endorsed measures requires hand abstraction of data from either paper charts or an electronic health record. Finally, most time stamps required for this abstraction would most likely be available in an electronic health record rendering this concern moot.
c. Dr. Phelan proposes an alternate scoring methodology and alternate bundle. We respectfully submit that this proposal is untested in terms of its evidence basis, reliability and validity. Modification of the composite measure to such an extent at this level of review would be an untested venture. At this time the only measure under consideration that could meet the required testing is that provided by the developer.
d. Finally, Dr. Phelan remarks that, "prior to accepting a measure like this we would like to see the data abstraction algorithm and testing of the data collection process to assess its data collection burden and feasibility." The measure developer notes that the data collection was possible in more than 200 hospitals participating in the Surviving Sepsis Campaign. The links to the tools used are provided in the measure submission. In any event, no well developed measure is easily abstracted with the average number of pages of details in current CMS specifications manuals exceeding 30. Please see http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228772433589 for examples.

Submitted by Ms. Carmella Bocchino, MBA, RN

While we are supportive of this measure, it may be difficult to implement due to difficulties in identifying numerator criteria, which is likely given the bundling nature of hospital claims.

Reply to Ms. Bocchino from the measure developers:

While we appreciate that bundled claims may represent difficulty in data gathering, we remain confident that the approach is not substantially different from those that payers presently use for data gathering. For example, CMS routinely collects information on quality performance for indicators such as heart failure, acute myocardial infarction, SCIP, and pneumonia care among others from such sources with detailed specifications for the numerator.

Submitted by Ms. Kathleen Szumanski, MSN, RN, NE-BC

ENA appreciates the opportunity to review this bundled measure. The measure stewards have appropriately responded to questions related to the measure. We support the committee decision to endorse the measure.

Reply to Szumanski from the measure developers:

We appreciate the commentary and support for this measure.

Submitted by Dr. Christopher Fee, MD, Society for Academic Emergency Medicine (SAEM)

Thank you for allowing the Society for Academic Emergency Medicine (SAEM) to comment on NQF Measure #0500: Severe sepsis and septic shock management bundle. Although SAEM supports the Surviving Sepsis Campaign (SSC) guidelines, we have significant concerns regarding the validity of this measure.

SAEM is concerned with the definition of the time of presentation. We support timely interventions for severe sepsis/septic shock but question the validity of triage time as time zero. Patients may present with uncomplicated infections or non-specific complaints only to develop severe sepsis/septic shock hours later. The current measure calls for bundle element compliance within three hours of triage (rather than from the time the patient meets severe sepsis/septic shock criteria). Thus, if a patient does not meet criteria until three or more hours after presenting to triage, even if all bundle elements are rapidly achieved, the hospital will be non-compliant despite providing outstanding, evidence-based care. Time-based measures that potentially start the clock ticking prior to patients meeting the defining criteria of the syndrome in question have to be recognized as invalid.

The developers responded that ED patients with infections are “somewhere on the natural trajectory of becoming septic regardless of point of presentation.” Statements such as this encourage overly aggressive treatment for patients who do not initially meet criteria for severe sepsis/septic shock due to provider concern of being deemed retrospectively “non-compliant” should the patients’ condition subsequently change. The developers state “if the patient who becomes hypotensive or has a high lactate does so in the ED, the reason for the presentation to the ED is severe sepsis or shock.” While this is true in cases where criteria are met at triage, it’s absolutely not the case for those who only do so hours later. Patients present with chief complaints (which are often non-specific), not diagnoses. Providers and hospitals should not be penalized for failing to meet invalid measures when the care they provide is in keeping with evidence-based guidelines.

SAEM also has concerns regarding recommendation F. While we recognize that the SSC recommends CVP monitoring (an unreliable and seldom

followed parameter), both it and measuring central venous oxygen saturation are only supported by one single center clinical trial (as such limited evidence supports its use). There are presently 3 ongoing large multinational clinical trials investigating if these measures are actually valid indicators of resuscitation adequacy. By definition, ongoing phase III trials suggest equipoise and strongly suggest against these quality parameters being recommended. In fact, the newest SSC guidelines (to be released in the coming months) acknowledge that not all centers of care have the same resources to allow these recommendations.

Reply to Fee from the measure developers:

We appreciate the opportunity to address the comments of Dr. Christopher Fee on behalf of the Society for Academic Emergency Medicine (SAEM). Dr. Fee's comments reflect similar concerns as his colleague's from ACEP and our responses will reference our previous answers for brevity where possible.

1. Dr. Fee's initial concern is that the time of presentation defined as triage time is a threat to validity since the chance exists that a patient may not develop signs and symptoms of severe sepsis until well after certain time based elements of the bundle have elapsed. He states that, "Time-based measures that potentially start the clock ticking prior to patients meeting the defining criteria of the syndrome in question have to be recognized as invalid."
    a. We respectfully submit that this concern reflects an unusual circumstance: that a patient coincidentally presented to the ED with some process other than severe sepsis and then developed severe sepsis within 3 hours (to use the cited time range). We sincerely believe this is not often the key issue. To illustrate a couple scenarios: those presenting with common diagnoses such as limb fractures do not spontaneously transform into severe sepsis cases; on the opposite end of the spectrum, those presenting with severe pneumonia and systemic symptoms were almost always certainly verging on severe sepsis and should have raised the index of suspicion substantially in the ED. We wish to suggest it is more often the case that a patient presents with signs and symptoms of severe sepsis that are poorly detected by inadequate screening methods in the emergency department.
    b. Although there may be an honest difference of belief, from an objective statistical standpoint, cases of the sort that Dr. Fee describes did not occur with sufficient frequency to disturb the calculations demonstrating validity in the submission. Even if such patients do present some percent of the time, they do not skew the data to suggest invalidity.
    c. We respectfully note that Dr. Fee's concern actually serves to point out that the variability in ED care is not so much the variation imposed by starting the clock at triage versus the recorded onset of

signs and symptoms of severe sepsis. The variation in ED care is the variation inherent in *a provider's clinical detection of severe sepsis*. We submit is not usually the patient who fails to manifest appropriate evidence of severe sepsis, rather it is *the poor systems that ED's presently use to detect severe sepsis* that leads to the discrepancy in times. Poor screening at triage, mis-triage, relegation to a lower level of care, delays in being seen by physician providers, and delays in laboratory care or turnaround time are the most common culprits *for failure to detect severe sepsis in a timely fashion* in the ED. Stated differently, the majority of cases that "change before our eyes" are not from the natural history of disease progression but due to poor detection and inadequate screening at presentation.

d. We respectfully suggest that Dr. Fee's statement that the strategy "encourage[s] overly aggressive treatment for patients who do not initially meet criteria for severe sepsis/septic shock due to provider concern of being deemed retrospectively 'non-compliant' should the patients' condition subsequently change" is actually not often the problem. As stated above, the patient is not usually the changing factor. Beyond that consideration, we respectfully submit that the data we have provided is a strong suggestion the measure does not result in "overly aggressive" medicine. The evidence provided in the submission actually shows that compliance with the composite measure improves over time and mortality decreases. The clinical reality might be different if mortality increased, but it does not.

e. Although we appreciate Dr. Fee's comments and perspective, we must respectfully conclude that he provides no firm evidence to demonstrate that the composite measurement strategy is invalid.

2. Dr. Fee's second concern mirrors Dr. Phelan's concern regarding element F in the numerator statement (*In the event of persistent arterial hypotension despite volume resuscitation (septic shock) or initial lactate >=4 mmol/L (36 mg/dl) measure CVP and central venous oxygen saturation*).

   a. Please see comments under Dr. Phelan's similar remarks in section 4b above for a proper refutation of these concerns.

   b. Dr. Fee comments that CVP monitoring is "an unreliable and seldom followed parameter." We respectfully ask the committee to note

that abandoning element F would provide clinicians no substitute parameter to assess intravascular volume in these patients. Even if Dr. Fee were to suggest a substitute indicator, we must point out that no other indicator has been demonstrated to be more accurate in this patient population. Nearly all experts agree on the centrality of adequate intravascular filling as a major concern for severely septic patient and that some estimate is required for proper care. The evidence basis assembled in the submission clearly shows a decreased mortality with adherence to the composite measure relied on use of CVP to make such an evaluation.

c. Dr. Fee points to "3 ongoing large multinational clinical trials investigating if these measures [CVP and $ScvO_2$] are actually valid indicators of resuscitation adequacy" and suggests that equipoise should prevent action. We respectfully submit that the statement is problematic for several reasons:
    i. The statement presumes that the purpose of the PROMISE, ARISE, and PROCESS trials is strictly to assess the adequacy of CVP and $ScvO_2$. We suggest that this is simply untrue. The effectiveness of these parameters is in fact only a secondary or tertiary analysis of each trial. Please see: https://www.promisetrial.org/, http://clinicaltrials.gov/show/NCT00975793, and https://crisma.upmc.com/processtrial/info2.asp.
    ii. Dr. Fee states that CVP and "measuring central venous oxygen saturation [$ScvO_2$] are only supported by one single center clinical trial (as such limited evidence supports its use)" however we respectfully note this is inaccurate. More than 40 papers in support of CVP and $ScvO_2$ were provided to the NQF maintenance committee in the original submission for evaluation. In review of these papers, the maintenance committee voted to endorse the scientific evidence basis for the recommended changes.
    iii. Please note that the EGDT study was not original and was based on 45 years of expert opinion in emergency medicine and critical care.[18] It is the second most cited publication in critical care and emergency medicine in the last 50 years[19] and its findings have been replicated more than 50 times in

the last 11 years in over 30,000 patients. We believe that to hold back a life saving therapy because of pending trials is a disservice to patient care. Two of the trials mentioned by Dr. Fee are being conducted outside the US and are incomplete. The US trial that began in 2008 is still incomplete. During this trial's run, mortality due to severe sepsis and septic shock has decreased 10-12% over the last decade, which suggests that the treatment effect may diminish the findings of these current randomized trials.[20] Although randomized controlled trials are considered the research standard, their use in evaluating therapies among the critically ill have been called into question for multiple reasons.[21] Multiple investigators have shown that estimates of treatment effects in large case-controlled observational studies are qualitatively similar to those obtained in randomized, controlled trials. Thus, the results of large observational studies such as GENESIS trial[5] provide important and equally acceptable contributions to the science of severe sepsis and septic shock. [22,23]

Submitted by ACEP QPC, Quality & Performance Committee; Submitted by Mr. Dainsworth Chambers

ACEP again thanks the Steering Committee for the opportunity to comment on the proposed changes to the NQF Endorsed Measure #0500. Our three primary concerns remain 1) the issues surrounding the reliability of triage being time zero, 2) the lack of evidence for the CVP measure component in the ED, and 3) the feasibility of abstracting the composite measure. ACEP has serious concerns surrounding the lack of evidence for measuring CVP as a surrogate for intravascular volume. NQF's "Composite Measure Evaluation Framework" clearly states in Table 1 that "the individual measures included in the composite or sub composite must be either NQF endorsed; or assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure."7Currently there are no NQF-endorsed measures that address CVP in septic patients. Although the measure developer submitted a number of quality indicators as part of a 6-hour ED severe sepsis bundle in 2007, several of the component indicators including CVP were not included in the currently NQF endorsed measure #0500, because they did not meet the NQF criteria for scientific acceptability as component measures at that time. Since then the measure developers have now cited five additional studies in which multivariate logistic regression demonstrated no independent effect on mortality in patients who achieve CVP targets versus patients who do not. (Castellanos-Ortega 2010, Nguyen 2007, Jeon 2012, Levy 2010, Cannon 2010).

Despite concerns raised by several Steering Committee members who noted CVP is one option of many potential surrogates for intravascular volume, they were not permitted to re-address their concerns regarding the validity of evidence when the appendix labeled "NQF Component Item Measure Analysis to Justify Inclusion in Composite" was provided at the end of the Steering Committee meeting. Measure developers also responded that only about 15 percent of patients end up needing a CVP when in compliance with the bundle, yet in the Surviving Sepsis Campaign quality improvement study 7,854 of the 15,022 patients in that study were eligible for the CVP indicator, and this number will vary from hospital to hospital based on case mix. NQF's "Composite Measure Evaluation Framework" clearly states that "all of the component measures must individually meet evaluation criteria," and this component does not meet the evidentiary threshold.

In conclusion, we respectfully request that the Steering Committee give serious re-consideration of the NQF criteria for validity, reliability, and feasibility for the recently proposed measure 0500: Severe Sepsis and Septic Shock Management Bundle. We urge the measure developers to work with the appropriate technical experts and stakeholders to address these questions.

Reply to Chambers from the measure developers:

Mr. Dainsworth Chambers submits additional comments which the measure developer is pleased to review and respond to here. Although Mr. Chambers reiterates his objections to the reliability of triage being time zero and the feasibility of abstracting the composite measure, these objections are more than sufficiently refuted in his initial comments. The feasibility of abstracting the composite measure is clearly established and demonstrated to be no more burdensome than existing data abstraction regimens in Dr. Phelan's comments above. The new critique that Mr. Chambers brings to bear in his second set of comments is around "the lack of evidence for the CVP measure component in the ED."

1. Mr. Chambers remarks that "ACEP has serious concerns surrounding the lack of evidence for measuring CVP as a surrogate for intravascular volume. NQF's 'Composite Measure Evaluation Framework' clearly states in Table 1 that 'the individual measures included in the composite or sub composite must be either NQF endorsed; or assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure.' Currently there are no NQF-endorsed measures that address CVP in septic patients."
    a. Because there is significant overlap in the questions and commentary to the measure, please see our responses to Dr. Phelan's comments in section 4b of our responses to his concerns. Specifically, we respectfully note that the next sentence (Table 1, page 5) in the Composite Measure Evaluation Framework after the above sentence goes on to state, "A component measure might not be important enough in its own right as an individual measure, but it could be determined to be an important component of a composite." We believe such is the case in this instance.

b. Mr. Chambers additionally notes that "the measure developers have now cited five additional studies in which multivariate logistic regression demonstrated no independent effect on mortality in patients who achieve CVP targets versus patients who do not. (Castellanos-Ortega 2010, Nguyen 2007, Jeon 2012, Levy 2010, Cannon 2010)."
    i. We wish to respectfully comment that the studies cited by Mr. Chambers used analyses that were not primary end points but rather secondary analyses relying on logistic regression to ascertain an "independent" effect on mortality. The suggestion that the composite indicator, which was tested in each of these papers, could be peeled apart through logistic regression when it was never tested as a primary end point to begin is unlikely.
    ii. Of note, many of these studies Mr. Chambers cites show reaching the $ScvO_2$ target which is obtained by CVP placement to be significantly associated with improved mortality. Thus, CVP and $ScvO_2$ are linked composites of cardiovascular physiology. In a meta-analysis of over 10,000 patients, Chamberlain et al showed that patients attaining the $ScvO_2$ target were twice as likely to survive as those who do not.[24]
    iii. We respectfully suggest, following the NQF's reasoning, that CVP as "[a] component measure might not be important enough in its own right as an individual measure, but it could be determined to be an important component of a composite." Moreover, CVP plays a key role in the composite measure in each of the trials cited by Mr. Chambers *permitting some assessment of intravascular volume from which the clinician may draw inferences about further care*. Contrary to Mr. Chambers suggestion, we believe clinicians should not be left to treat severely septic patients with no method of assessing intravascular volume. We respectfully note that no trial on severely septic patients has endorsed this strategy. In this regard, it would seem to require a high evidentiary bar to intentionally omit this standard from the composite

measure even though it may not stand alone as an individual component measure. We believe such a circumstance is the purpose of the NQF rule that a composite may include an indicator that would not necessarily stand alone (Composite Measure Evaluation Framework ,Table 1, page 5).

c. Mr. Chambers notes that "several Steering Committee members who noted CVP is one option of many potential surrogates for intravascular volume."
    i. We must respectfully remark that this is not sufficient proof that these other surrogates are better evidence of intravascular volume in this patient population. Indeed, all of the trials including those that Mr. Chambers has cited above (Castellanos-Ortega 2010, Nguyen 2007, Jeon 2012, Levy 2010, Cannon 2010) used CVP as the only estimate for intravascular volume.
    ii. We believe Mr. Chambers' concern should be tempered insofar as nothing in the composite measure limits providers from using other measures of intravascular volume in addition to CVP.
    iii. We respectfully suggest that Mr. Chamber's reference to papers that used CVP underscores the consensus assessment among experts that some measure of intravascular volume assessment is necessary in treating severely septic patients. We believe it should not be lost on the committee that these experts chose CVP just as we chosen.
    iv. Finally, the committee evaluating the evidence during the maintenance cycle fully considered all of the available evidence in support of and against the proposed changes. The committee concluded that the proposed changes meet the required scientific evidentiary standards. The submission included sub analyses of the component measures with respect to reliability and validity as well. In each instance, the committee approved the composite.

d. Lastly, Mr. Chamber's urges that, "…the measure developers [] work with the appropriate technical experts and stakeholders to

address these questions." The measure developers very respectfully and humbly suggest that we are appropriate content experts to make these recommendations. We believe we have respectfully engaged our colleagues both through the NQF process and outside the process to discuss these matters. We are pleased to continue the dialogue, but we believe that the evidence is presently sufficiently mature to make adopt the composite measure based on the maintenance committee's endorsement of importance, sufficiency of evidence, reliability, validity and feasibility.

Submitted by Mr. Reginald Lavender

Edwards' supports the continued endorsement of measure #0500: Severe Sepsis and Septic Shock: Management Bundle for endorsement. Sepsis is a devastating condition that has considerable impact to patients, hospitals and the overall health system. There are an estimated 1.14 million cases of sepsis or septicemia in the U.S. every year.[i] With mortality rate in excess of 28%,[ii]over 300,000 of these patients die annually, which is greater than the number of deaths due to breast, prostate, and lung cancer combined.[iii]

Even with continued endorsement of this measure, a clear gap remains in the need for more endorsed measures addressing sepsis, septic shock and septic management. This provides public and private payers with an opportunity to adopt existing, validated and tested, measures into incentive-based reporting programs to improve the care of patients with sepsis. Highmark's use of this measure in its pay for performance programs for the past two years highlights the need for measures addressing this high-burden condition.[iv]

In 2002, the Surviving Sepsis (SS) Campaign designed the bundle to allow teams to follow the timing, sequence, and goals of the bundle and to achieve a 25 percent reduction in mortality from sepsis or septic shock. The steering committee questioned whether the sepsis quality measure addressing a bundle should be endorsed versus specific validated elements of the bundle. The SS Campaign noted that by making the bundles standard practice, there is elimination of piecemeal or chaotically applied standards for sepsis care that exist in many clinical environments today. The sepsis bundle has also been rigorously used and tested in The Surviving Sepsis Campaign. This measure has proven success of improved outcomes among one hundred sixty-five hospitals around the world that voluntarily collected data to demonstrate the beneficial effects. Compliance with the sepsis bundles increased by 20 percent over two years and was associated with 7 percent reduction in hospital mortality.[v]

Finally, Edwards supports the continued endorsement of this measure in alignment with CMS' efforts to adopt NQF-endorsed and MAP-recommended measures into Medicare public reporting programs

[i]Hall MJ, et al. Inpatient care for septicemia or sepsis: a challenge for patients and hospitals. NCHS data brief, no 62. Hyattsville, MD: National Center for Health Statistics. 2011.[ii]Angus DC, et al. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Crit Care Med 2001;29:1303-1310.[iii]Siegel R, et al. Cancer Statistics, 2012. CA Cancer J Clin 2012;62:10-29.[iv]National Quality Forum. National Voluntary Consensus Standards: Infectious Disease Endorsement Maintenance 2012, Addendum Report. October 22, 2012. [v]Surviving Sepsis Campaign. Available at: http://www.survivingsepsis.org/Pages/default.aspx.

Reply to Lavender from the measure developers:

We appreciate the commentary and support for this measure.

Submitted by Ron Elkin, MD

Measures presently supported by NQF for severe sepsis and septic shock, without targets for CVP and ScvO2, will not result in optimal outcomes for this common and lethal disorder. Guidance is necessary for management of CVP and $ScvO_2$. In the absence of targets for CVP and $ScvO_2$ in the original Rivers study, there would be no "goal- directed" therapy, no mortality benefit, no Surviving Sepsis Campaign, and no NQF proposal to address severe sepsis and septic shock.

Advocacy for early goal-directed therapy requires acceptance of only 2 simple, indisputable principles that most consider well documented and self-evident:

1) Tissue injury, morbidity, and mortality due to the hypoperfusion of severe sepsis and septic shock are time sensitive, similar to that occurring with myocardial infarction and stroke. Outcome will improve with early correction of hypoperfusion.

2) Blood pressure, other vital signs, and bedside examination are inaccurate indicators of the adequacy of tissue perfusion.

The landmark study by Rivers et al therefore supplemented conventional CVP and BP resuscitation targets with an indirect but useful surrogate for tissue perfusion – $ScvO_2$. As a consequence of the $ScvO_2$ target for resuscitation, a dramatic reduction in mortality was demonstrated, confirmed by numerous subsequent trials with historical controls.

There are alternatives to $ScvO_2$ for estimating the adequacy of tissue perfusion, but none are universally available, applicable or sufficiently reliable, and none have been rigorously compared to $ScvO_2$. Lactate clearance, with several caveats, should be regarded as a valuable supplement, rather than a substitute for $ScvO_2$.

It is an error to attempt correlation of outcome with single goals of resuscitation. The goals must be examined together rather than in isolation. For example, despite the CVP oft-maligned as poorly correlating with blood volume, fluid responsiveness, or outcome, few clinicians would forego intervention for a low CVP in the presence of hypotension, vasopressors, or hypoperfusion. As a second example, an improving $ScvO_2$ with fluid

administration is indeed evidence of “fluid responsiveness”. As crude as these current goals may be, they are widely available and collectively useful in reducing mortality.

At this time, it is imprudent to ignore the specific targets used in early goal-directed therapy as studied by Rivers. Failure to specify targets does a disservice to approximately one million patients per year suffering from these lethal disorders. Until evidence for better methodologies with superior outcomes emerges, I respectfully request adoption by NQF of the resuscitation targets validated by Rivers and recommended by the Surviving Sepsis Campaign.

Ron Elkin, M.D.

California Pacific Medical Center

San Francisco

Reply to Elkin from the measure developers:

We appreciate Dr. Elkin’s comments regarding targets for CVP and $ScvO_2$ measurement. The measure developers concur that the purpose in measuring the CVP and $ScvO_2$ as process measures in the composite measure is to take appropriate action. To that end, the appropriate actions are listed in the 2008 and 2012 Surviving Sepsis Campaign Guidelines. In addition, the appropriate actions in all of the trials cited in support of this measure were to optimize these parameters to CVP >= 8 and $ScvO_2$ >= 70%. Since the composite is entirely a process measure, these goals were left out of the text. We suggest that an asterisk be placed after element F in the numerator statement for reference as follows:

> *In the event of persistent arterial hypotension despite volume resuscitation (septic shock) or initial lactate >=4 mmol/L (36 mg/dl) measure CVP and central venous oxygen saturation.**
>
> **Targets for quantitative resuscitation included in the 2012 Surviving Sepsis Guidelines are CVP of ≥8 mm Hg and ScvO2 of ≥70%.*

References

1. Reed K, McBratney S, Thompson D, Nicholas C. Healthgrades emergency medicine in american hospitals study. Health Grades 2011;The First Annual Report:1-28.

2. Carr GE, Yuen TC, McConville JF, et al. Early Cardiac Arrest in Patients Hospitalized With Pneumonia: A Report From the American Heart Association's Get With the Guidelines-Resuscitation Program. Chest 2012;141:1528-36.

3. Wang HE, Shapiro NI, Angus DC, Yealy DM. National estimates of severe sepsis in United States emergency departments. Crit Care Med 2007;35:1928-36.

4. Andrews R, Elixhauser A. The national hospital bill: growth trends and 2005 update on the most expensive conditions by payer. Healthcare Cost and Utilization Project 2007;http://www.hcup-us.ahrq.gov/reports/statbriefs/sb42.pdf:1-13.

5. Cannon CM, Holthaus CV, Zubrow MT, et al. The GENESIS Project (GENeralized Early Sepsis Intervention Strategies): A Multicenter Quality Improvement Collaborative. J Intensive Care Med 2012 (published ahead of print, 8/21/2012).

6. Marchick MR, Kline JA, Jones AE. The significance of non-sustained hypotension in emergency department patients with sepsis. Intensive Care Med 2009;35:1261-4.

7. Mikkelsen ME, Miltiades AN, Gaieski DF, et al. Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Crit Care Med 2009;37:1670-7.

8. Trzeciak S, Dellinger RP, Chansky ME, et al. Serum lactate as a predictor of mortality in patients with infection. Intensive Care Med 2007;33:970-7.

9. Shapiro NI, Howell MD, Talmor D, et al. Serum lactate as a predictor of mortality in emergency department patients with infection. Ann Emerg Med 2005;45:524-8.

10. Pearse RM. Extending the role of lactate measurement into the prehospital environment. Crit Care 2009;13:115.

11. Yang CS, Qiu HB, Huang YZ, et al. [Prospective research on the prognosis of septic shock based on the change of lactate concentration in arterial blood]. Zhonghua Wai Ke Za Zhi 2009;47:685-8.

12. Rivers E, Nguyen B, Havstad S, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 2001;345:1368-77.

13. Levy MM, Dellinger RP, Townsend SR, et al. The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. Critical Care Medicine 2010;38:367-74.

14. Kashiouris M, Kashyap R, Jaffer-Sathick I, Gajic O, Cartin-Ceba R. Association between delays in central venous catherter utilization and organ dysfunction in patients with severe sepsis and septic shock. Critical Care Medicine 2011;39(12):142.

15. Varpula M, Tallgren M, Saukkonen K, Voipio-Pulkki LM, Pettila V. Hemodynamic variables related to outcome in septic shock. Intensive Care Med 2005;31:1066-71.

16. Walkey AJ. Utilization Patterns and Outcomes Associated with the Central Venous Catheter in Septic Shock: A Population-Based Study. Crit Care Med 2012 (accepted for publication).

17. Kumar A, Ellis P, Arabi Y, et al. Initiation of Inappropriate Antimicrobial Therapy Results in a 5-Fold Reduction of Survival in Human Septic Shock. Chest 2009.

18. Wilson RF, Wilson JA, Gibson D, Sibbald WJ. Shock in the emergency department. Jacep 1976;5:678-90.

19. Tao T, Zhao X, Lou J, et al. The top cited clinical research articles on sepsis: a bibliometric analysis. Crit Care 2012;16:R110.

20. Kumar G, Kumar N, Taneja A, et al. Nationwide trends of severe sepsis in the 21st century (2000-2007). Chest 2011;140:1223-31.

21. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. Crit Care Med 2010;38:S534-8.

22. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med 2000;342:1878-86.

23. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000;342:1887-92.

24. Chamberlain DJ, Willis EM, Bersten AB. The severe sepsis bundles as processes of care: A meta-analysis. Aust Crit Care 2011.