Sutter Health CPMC
We Plus You

Sean R. Townsend, MD
Vice President Quality & Safety

Stern Building
2330 Clay St., #301
(415) 600-5770 Phone
(415) 600-1541 Fax
San Francisco, CA 94115

Via Electronic Submission (www.qualityforum.org)

June 25, 2014

Consensus Standards Approval Committee (CSAC)
National Quality Forum
1030 15th Street NW
Suite 800
Washington DC 20005

Dear Committee Members:

Executive Summary:

- An ad hoc committee of NQF was charged with reviewing the quantity, quality, and consistency (QQC) of measure #0500 after publication of the Protocolized Care for Early Septic Shock (ProCESS) trial.
- ProCESS concluded that a certain element of #0500 (specifically element F) was not harmful, but only as effective as some other strategies employed in caring for septic shock patients.
- The ad hoc committee considered questions beyond the findings in ProCESS based on speculation and fear without citation to evidence base to support their concerns. The evidence basis actually shows central lines are less harmful than peripheral lines for critically ill patients.
- The ad hoc committee was unprepared by NQF staff for the rigor of the evaluation.
- The ad hoc committee did not consider the best yardstick of the QQC of the evidence, a meta-analysis that plainly shows the QQC is in favor of element F.
- The ad hoc committee was swept up by the "newness" of the trial and the associated buzz of the findings, but unfamiliar with the entire evidence basis. The committee should have been better informed by NQF staff about the evidence basis and to resist the temptation to judge new evidence outside of the old, a fatal error in conduct for a scientific review panel.
- Sepsis measures are a Measures Application Partnership (MAP) and Department of Health and Human Services priority. CMS has now thusly proposed NQF #0500 in the FY 2015 Inpatient Prospective Payment System Proposed Rule, File Code CMS-1607-P for implementation in 2017. This uncertain process undertaken at NQF disrupts progress on a national priority.
- The technical actions of the ad hoc committee were fatally flawed from a scientific acceptability standpoint, standing in direct contrast to NQF's publication standards for composite measures, and must compel CSAC to reject the ad hoc committee's approach.
- The best approach is to return this question to committee for determination after the evidence base is complete within 1 year with 2 new pending trials (the Australian Resuscitation In Sepsis Evaluation Randomised Controlled Trial [ARISE] and The Protocolised Management in Sepsis Trial [ProMISe]). At that time the committee will need to determine whether to reject the entire composite or continue as written because there is no precedent to remove specific elements of a composite without conducting necessary reliability and validity testing of the remaining items.
- Finally, because ProCESS showed no harm from items in element F (but *by itself* showed no benefit either – other studies disagree), rejecting #0500 now when demonstrated to be effective

though the Consensus Development Process would be inappropriate and #0500 should remain endorsed.

Introduction & Background:

As a key strategic partner to Dr. Emanuel Rivers, the measure steward for NQF #0500, and trusted consultant on its development, I am writing in opposition to the changes approved on a 7-11 basis by an ad hoc review committee formed under NQF's Patient Safety Committee to review the Protocolized Care for Early Septic Shock (ProCESS) trial, which was published on March 18, 2014.

The charge of the ad hoc review committee was to review the evidence supporting item "F" of measure #0500, which is the element related to central venous pressure (CVP) and central venous oxygen saturation ($ScvO_2$) monitoring via a central venous catheter monitoring for patients with septic shock.

The committee was asked to make an assessment as to whether the quantity, quality, and consistency (QQC) of the evidence base as a whole supports item "F" of the bundle.

Inadequacy of the Preparedness of the Committee for the Scientific Question:

ProCESS concluded that there were no detectable mortality differences in patients receiving Early Goal Directed Therapy (EGDT), essentially element F of the composite measure, compared to what the trial termed "usual care." This conclusion, which remains highly suspect due to contamination of intervention elements (EGDT) in the control group ("usual care") from a single trial, even if the conclusion was accurate, had exactly zero effect on the QQC of the evidence. How can we be sure? Please see Dr. Rivers' meta-analysis in his letter of opposition, which graphically demonstrates that at trial showing *no difference between 2 therapies* (neither is better or worse) has no effect on the central tendency of the data when the data already falls substantially in favor of EGDT for mortality reduction. We know for a fact that in tens of thousands of patients who have received EGDT, mortality versus standard care patterns is reduced.

This alone should have ended the debate about the quantity (10's of thousands), quality (3 RCT's 30+ observational trials), consistency (no trials showing harm, almost all showing benefit), but the committee of well intended individuals, opined fearfully about scientifically unrelated questions.

The transcript shows record of committee members meandering in their conversation about types of lines that didn't apply to EGDT (Swan-Ganz catheters, day 1 page 514), discussing "scarred down" necks (day 1, April 17 2014, page 531), and the alleged harms from central line placement ("[w]e do not need to mandate a central line right now, especially when waiting a year or two for those other studies to come out leads to possibly more unintentional complications, mortality associated with lines if patients don't need it" (day 1, page 532). There were comments about "whizbang technology" not being as good as the old stuff (day 1, page 516).

The Facts about the Fear:

ProCESS itself showed *zero differences* in line associated complications. Dr. Yealy, the study's lead author, was forced to confirm this fact on day 2 April 22, 2014, page 39 of the "Patient Safety SC_Post Meeting Call_042214_Transcript.doc," "[m]andating one approach when in fact two other approaches are equally good exposes patients to intervention that they may not benefit from and that they could have harm from. **Obviously, we did not detect that on the trial**…" (emphasis supplied). Moreover, in a disease process (septic shock) where there is at least a 1 in 5 chance of dying, the established benefit far outweighs the risk of death from a central line, estimated to be 1 in 1000 (1,2). We now also know that *central lines are less harmful for critically ill patients than peripherally placed lines for early resuscitation* from a multicenter randomized controlled trial – suggesting just the opposite of the committee's speculative fears – **perhipheral lines result in more complications** (3).

The committee failed to mention the provided meta-analysis entirely although it was provided. There is no discussion of the import of the meta-analysis on the record, (day 1, day 2). The meta-analysis is itself plainly a review of the quality quantity and consistency of the evidence.

The committee had no detectable familiarity with the other trials in the meta-analysis aside from the original Rivers' trial.

One committee member remarked in a moment of candor, "[a]nd so what are we trying to accomplish here that the experts themselves are never going to be satisfied with? (Laughter.)" (day 1, page 534).

The committee therefore failed to evaluate the QQC of the evidence in a scientifically rigorous fashion in favor of being blindsided by a shiny new object in their view. This was reinforced by a degree of associated media hype generated in part due to publication in the nation's pre-eminent medical journal (where incidentally the Rivers' trial was also published, albeit 13 years prior). The scene reflected and felt very much like an episode of "The View," rather than a scientific exercise. It was truly a miscarriage of the committee's responsibility, although well-intended and with very intelligent committee members from highly regarded backgrounds. NQF staff had a responsibility to better prepare the esteemed members.

Importance of this Measure:

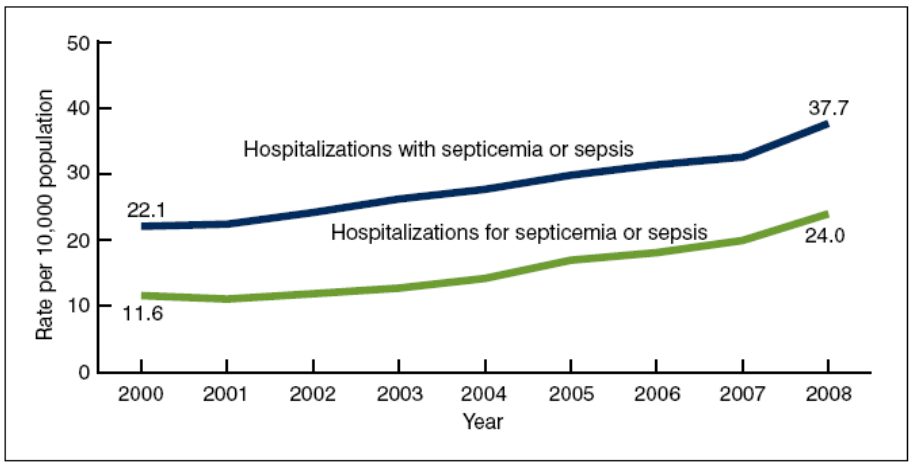
A committee empowered to review the QQC of the evidence had better know the evidence well and make well thought out recommendations when the Measures Application Partnership (MAP) has ranked severe sepsis and septic shock a prime area of interest for appropriate and effective measures. MAP was created to provide input to the Department of Health and Human Services (HHS) on the selection of performance measures for public reporting and performance-based payment programs. NQF was selected by HHS to fulfill a statutory requirement to convene multi-stakeholder groups to:

- Identify the best available performance measures for use in specific applications.
- Provide input to HHS on measures for use in public reporting, performance-based payment, and other programs.
- Encourage alignment of public- and private-sector performance measurement efforts.

HHS is particularly interested in the disease and getting the measure properly drafted because the incidence of septicemia has nearly doubled in the last decade.

Hospitalization rates for septicemia or sepsis more than doubled from 2000 through 2008.

Figure 1. Hospitalizations for and with septicemia or sepsis



NOTE: Significant linear trend from 2000 through 2008 for both categories.
SOURCE: CDC/NCHS, National Hospital Discharge Survey, 2000–2008.

For an NQF committee to alter NQF #0500 in light of a trial that showed no difference in mortality, that is *no harm* to applying the consensus measure, after finally having approved it unprecedented. It is destructive as a matter of policy making as well since the measure now sits as part of the 2017 inpatient quality reporting (IQR) proposed rule that HHS issued this year (see FY 2015 Inpatient Prospective Payment System Proposed Rule, File Code CMS-1607-P).

Technical Inadequacy of the ad hoc Committee's Actions:

The ad hoc committee was wrongly empowered by NQF staff to examine an element of the composite measure #0500 and permit its removal. This charge stands in conflict with the definition of a composite measure and the requirements of approval for a composite measure through the Consensus Development Process. These standards are enumerated in the NQF publication, "Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures" published October 2012.

A composite measure, unlike an individual measure, "is a combination of two or more individual measures in a single measure that results in a single score." (see page 5). In developing a composite measure, NQF requires that the final measure be tested -- "[f]inally, as with all measures, test the composite to determine if it is a reliable and valid indicator of quality healthcare," (see page 4). NQF proceeded to dismantle composite without requiring the appropriate testing of the newly devised measure the ad hoc committee unilaterally created. They were granted some blanket permission not to re-test the re-designed composite for features of scientific acceptability.

In particular, the ad hoc committee shorn from the composite measure element F, yet assumed that the measure as a whole retained scientifically acceptable measure properties. The original composite, including item F, was shown to meet reliability testing of the composite proving that "the results were repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period" (Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures, page 11). This reliability testing was accomplished for the composite measure using RAND methodology at the direction of NQF's methodologist *and never included testing without element F*.

Although the Composite Measure Evaluation Framework suggests that the elements of composite measures must meet the standards of scientific acceptability for individual measures, the framework also demands that the composite be fully and independently tested in this regard ("[e]ven though all of the component measures must individually meet evaluation criteria, the composite measure as a whole also must meet evaluation criteria," page 5). Thus, the shell left behind by the committee's action cannot be said to have statistically proven reliability a specified requirement in the above publication for a composite measure. Thus, this rationale that the independent elements are all valid is no defense to permit a committee to unilaterally draft a new composite without completing the appropriate testing.

The same basic idea that reliability testing is key to the scientific acceptability of a measure applies to validity. Specifically, NQF's "Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures" requires at page 12 that, "[v]alidity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed." Again, the testing applied to #0500 was suggested and vetted by NQF's methodologist and statistically demonstrated in the original submission. The ad hoc committee's action to strip element F from the composite paid no heed to the impact of removing item F as regards validity testing – and nobody can scientifically say the new measure is valid. Certainly, the record reflects that ad hoc committee was not charged to even "systematically assess" the face validity of the shell left behind, the lower standard.

Moreover, the phrasing "the same results a high proportion of the time when assessed in the same population in the same time period" is not meaningless (Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures, page 11). The

data that supported reliability and validity testing of NQF #0500 came from 200 community hospitals in the United States between 2007-2009. The ProCESS trial was conducted in 31 academic medical centers between 2007-2014. Neither the setting nor the time periods are commensurate and to assume that the data based on a different setting over a longer time period are somehow substitutable into the community hospital setting over a different time period is a radical leap – and never would have passed muster in the original approval of the NQF #0500 for scientific acceptability. It should not do so now.

In short, the ad hoc committee was not properly charged with the task of considering how removing element F would affect the scientific acceptability of NQF #0500. **The ad hoc committee in reality had only one choice: leave the measure intact or toss the whole thing aside.** It was and remains not appropriate to toss the entire measure aside both because of the national imperative to have sepsis quality measures indentified by HHS, MAP and NQF. However, the next most expedient action, to just strip element F out of the composite, is also not appropriate because it is counter to NQF's own standards for scientific acceptability.

The action thus endorsed by the committee is unfortunately is not a permissible approach. It is also a step too far and entirely unnecessary when the ProCESS trial concluded that the items in element F were, in fact, still efficacious. Other strategies may also be efficacious, but there is no imperative to have moved in the only scientifically acceptable direction -- a decision to toss out the entire measure because adequate testing for scientific acceptability minus element F cannot be conducted. Had the committee been properly charged with the scientific ramifications of removing element F, i.e. the need to re-test the entire newly created composite and had the committee been educated on the absence of any data on which to do the required validity and reliability testing, **it is unlikely the committee would have voted to remove the entire measure where no harm was demonstrated.** Again, the committee was not properly instructed in its options and the scientific folly of removing elements of a composite measure in a wanton fashion.

Danger and Patient Harm in a Potential CSAC Endorsement the ad hoc Committee's Approach:

The newly minted composite measure that the ad hoc committee created is not only untested but immediately runs into the potential for great harm. The new composite requires a fluid bolus of 30 ml/kg for hypotension or lactate greater than 4 mmol/L (element D). If the patient remains hypotensive after that, the newly minted composite instructs the provider to place the patient on vasopressors (element E). No further assessment of the adequacy of volume status is required. No further assessment of the adequacy of tissue perfusion is required.

Essentially, the ad hoc committee's new composite measure relegates patients who did not get enough fluid administration to instead receive possibly ever increasing doses of vasopressors. The obvious harm and problem here is that such a strategy is untested in terms of mortality reduction and may cause increased mortality. Tissue ischemia and loss of digits and extremities is a known complication of applying vasopressors to under resuscitated patients (4). While this may seem unthinkable, this is precisely the state of affairs that existed prior to the resuscitation strategies advocated by the Surviving Sepsis Campaign – patients languished on vasopressors while under-resuscitated (5).

Nobody knows what will happen now with the ad hoc committee's untested composite because nobody has ever tested the effect of such measurement on provider's behavior using the ad hoc committee's approach.

While the ad hoc committee had hoped that providers would still have the option of assessing these items, but that it would not be mandatory to do so, nobody has ever tested what happens *when you measure providers* under such a strategy. The ProCESS trial was not a test of measurement strategies on providers. In particular, ProCESS did not instruct providers via a proscribed measurement strategy endorsed by regulators to give an initial fluid bolus and then put the patient on vasopressors as the only measurement strategy.

ProCESS was not a test of measure efficacy, it was a scientific trial. Only the data submitted in favor of NQF #0500 was a test of the effect of real world measures on patient outcomes. NQF has inadvertently set into motion an untested set of measures on provider's behavior, the effect of which is unknown.

Recommended Course of Action for CSAC:

Based on the foregoing, it is plain that for scientific integrity purposes, CSAC cannot ratify the actions of the ad hoc committee without entering into substantial scientific folly as regards scientific acceptability testing. NQF cannot be responsible for a strategy that is untested and may induce patient harm.

CSAC must therefore send this back to committee for further consideration. The proper course of action at this time, since ProCESS showed no harm to element F, is to leave the measure standing as a necessary advance in scientific care of severely septic and septic shock patients.

CSAC should consider charging the ad hoc committee with awaiting the 2 further trials to be published in the next year that will endeavor to answer the same questions as ProCESS. At that time the entire evidence base on element F (EGDT) will be complete. If the consensus of the trials is that element F is improper, the measure developer will submit new data for testing scientific acceptability without these items as part of a new composite once the real world testing is completed and gathered.

CSAC should not charge the ad hoc committee with suspending the entire measure as required for proper scientific acceptability testing since ProCESS did not demonstrate any harm to patients associated with the measure. Sepsis measures remain a national priority. Dismantling a good measure proven to be effective in the real world as a measurement strategy in the face of no demonstrated harm is not defensible as a matter of science or public policy.

The advantages to this approach are:

1. Preservation of national sepsis measures
2. No demonstrated harm to patients from continued measure endorsement
3. Preservation of scientific integrity of #0500 without creating a new measure untested for reliability and validity via removal of element F
4. No interference with HHS's proposed rule, FY 2015 Inpatient Prospective Payment System Proposed Rule, File Code CMS-1607-P
5. The advantage of viewing a complete evidence base before changes are made requiring yet further revisions to #0500.

Thank you for your consideration of these important matters.

Sincerely,

Sean R. Townsend, MD

References:

1. McGee DC, Gould MK. Preventing complications of central venous catheterization. N Engl J Med. 2003;348(12):1123.

2. Eisen LA, Narasimhan M, Berger JS, Mayo PH, Rosen MJ, Schneider RF. Mechanical complications of central venous catheters. J Intensive Care Med. 2006;21(1):40.
3. Ricard JD, Salomon L, Boyer A, et al. Central or peripheral catheters for initial venous access of ICU patients: a randomized controlled trial. Crit Care Med 2013;41:2108-15.
4. Beale RJ, Hollenberg SM, Vincent JL, Parrillo JE. Vasopressor and inotropic support in septic shock: an evidence-based review. Crit Care Med. 2004 Nov;32(11 Suppl):S455-65. Review.
5. 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.