NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Surgery Standing Committee—Measure Evaluation Web Meeting

The National Quality Forum (NQF) convened the Surgery Standing Committee for a web meeting on February 19, 2020 at the NQF offices in Washington, DC to evaluate one measure.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF staff reviewed the meeting objectives. Committee members each introduced themselves and disclosed any conflicts of interest. Standing Committee member Frederick Grover, MD was recused from discussion and voting for measure 0696. Alex Sox-Harris PhD, MS was recused from voting on Scientific Acceptability on 0696 due to his involvement on the Scientific Methods Panel. Quorum was met and maintained throughout the web meeting. One panel member joined the meeting late and two had to leave early. The vote totals reflect the members present and eligible for each vote.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and reviewed the Consensus Development Process (CDP) and the measure evaluation criteria. NQF staff provided a brief overview of the Scientific Method Panel (SMP) process, including the SMP's deliberations on measure 3537 *Intraoperative Hypotension among Non-Emergent Noncardiac Surgical Cases*. The SMP passed this measure on reliability (H-2; M-3; L-1; I-0) but gave it a low rating for validity (H-0; M-2; L-4; I-0). During SMP in-person meeting on October 28, 2019, the Panel determined that a variable included in the risk-adjustment model (length of surgery) could be directly affected by the presence of intraoperative hypotension (the subject of the measure). The SMP felt this was a threat to the validity of the measure results and the SMP informed the Standing Committee the measure was not eligible for Committee vote on validity.

Measure Evaluation

During the meeting, the Surgery Standing Committee evaluated one measure for endorsement consideration. The maintenance measure was recommended for continued endorsement. A summary of the Committee deliberations will be compiled and provided in the draft technical report. NQF will post the draft technical report on March 30, 2020 for public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

Rating Scale: H – High; M – Medium; L – Low; I – Insufficient; NA – Not Applicable

0696 STS CABG Composite Score (Society of Thoracic Surgeons)

Measure Steward/Developer Representatives at the Meeting

Jeff Jacobs, Mark Antman

Standing Committee Votes

- Evidence: Pass-14; No Pass-0
- Performance Gap: H-1; M-10; L-3; I-0

- Composite - Quality Construct and Rationale: H-5; M-9; L-0; I-0
- Reliability: Yes-14; No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's ratings for Reliability: H-0; M-6; L-1; I-0
- Validity: Yes-14; No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's ratings for Validity: H-2; M-4; L-0; I-0
- Composite Construction: H-6; M-8; L-0; I-0
- Feasibility: H-10; M-3; L-0; I-0
- Use: Pass-13; No Pass-0
- Usability: H-4; M-9; L-0; I-0

Standing Committee Recommendation for Endorsement: Yes-13; No-0

The Standing Committee recommended the measure for continued endorsement. Measure 0696 *Society of Thoracic Surgeons* (*STS) CABG Composite Score* consists of four domains comprised of 11 individual NQF-endorsed cardiac surgery measures. The Committee noted that evidence for the measure has not changed since its previous evaluation. The Committee discussed whether the measure was "topped out" with little room for improvement. Committee members noted that although the performance distribution among facilities appeared to be narrow, there was both meaningful opportunity for improvement overall and significant variability across STS participants in at least three of the four components (mortality, morbidity and medications) within the composite measure. The Committee agreed that the data on disparities were compelling across the individual domains, with increased risk for morbidity and mortality demonstrated for the female sex and African American race. When discussing the overall quality construct of the composite, the Committee noted that the components with the heaviest weighting in the composite had the least variation and those with the lower weighting had higher variation. The Committee wondered if this contributed to flattening out the distribution of overall performance. The developer responded by noting that the weighting scheme is supported by published consensus statements from an expert panel. In addition, the developer has found the weighting to have face validity with patient's considerations of the relative importance of the components. The Committee agreed that the measure meets NQF's criteria for composite quality construct.

The SMP evaluated the reliability and validity of 0696, rating both as "Moderate." The Standing Committee voted unanimously to accept the ratings of the SMP. Committee members noted that the content validity assessment demonstrated that there were significant and clinically meaningful differences in all four domains between hospitals with one- and three-star ratings. When discussing validity, Committee members questioned the face validity of star ratings to patients and if they aligned with consumer's understanding of what constitutes a one-star or three-star facility. The measure developer clarified that the star rating was designed to evaluate surgical programs and that the definition of the star rating and meaning behind it is available to the public on the STS web site. The

Committee and developers discussed the challenges of converting measure scores to star ratings. The Committee noted that the size of the confidence intervals could influence the star rating assigned. The developer acknowledged this and added that there is no perfect method for assigning star ratings but that they are using a well-accepted and tested methodology. The Committee concurred with this. The Committee discussed the composite construction including how the STS expert panel created the relative weighting scheme to assign the final composite score (and star ratings) and whether the composite score is meant to replace the individual four domain scores, or simply be used as a summary assessment. Committee members agreed that more granular assessments should still be available to STS participants to decide where to prioritize their quality improvement efforts and to patients so they can weigh what's important to them in choosing a hospital. The Committee agreed that the component measures fit the quality construct and that the weighting rules are in alignment with expert assessment and empirical testing.

The Committee noted that the data for this measure are collected as part of the STS Adult Cardiac Surgery database and had no major concerns regarding feasibility. When discussing the use subcriterion, the Committee expressed that additional information for consumers might be useful on the STS web site. The Committee noted that other STS public reporting programs have information on the volume of patients in the program and outcomes stratified by patient complexity. The Committee had no concerns with the usability of this measure.

Discussion of Related and Competing Measures

NQF staff noted that there are currently endorsed measures that are related to measure 0696. Staff facilitated a Committee discussion of whether the related measures were harmonized to the extent possible and if there is justification for endorsing multiple related measures. The first group of measures have the same measure focus as 0696, but different target populations:

- 2561 STS Aortic Valve Replacement (AVR) Composite Score
- 2563 STS Aortic Valve Replacement (AVR) + CABG Composite Score
- 3031 Mitral Valve Repair/Replacement (MVRR) Composite
- 3032 Mitral Valve Repair/Replacement (MVRR) + CABG Composite

The developer noted that they are currently working on an overall composite that would include all of these target populations; however, they stated there is still a need for procedure-specific measures as well. The Committee wondered if splitting the measures out into different procedures affects the ability to make meaningful quality observations. The developer noted this is handled by adjusting the measurement period so that each measure has enough patients included to calculate meaningful results. The developer and Committee agreed that it is important to have multiple ways of assessing and viewing quality and that quality may vary by type of operation. Both felt this is important information for providers and patients. The Committee raised the question of burden related to multiple measures and the developer noted that providers that participate in the registry must enter all relevant cases, so the information needed for the measures is captured regardless of measure calculation. The Committee was satisfied there was a justification for the related measures and that burden was minimized.

The second group of measures discussed have the same target population as 0696, but different foci:

- 0114 Risk-Adjusted Postoperative Renal Failure
- 0115 Risk-Adjusted Surgical Re-exploration
- 0116 Anti-Platelet Medication at Discharge
- 0117 Beta Blockade at Discharge
- 0118 Anti-Lipid Treatment Discharge

- 0119 Risk-Adjusted Operative Mortality for CABG
- 0127 Preoperative Beta Blockade
- 0129 Risk-Adjusted Postoperative Prolonged Intubation (Ventilation)
- 0130 Risk-Adjusted Deep Sternal Wound Infection
- 0131 Risk-Adjusted Stroke/Cerebrovascular Accident
- 0134 Use of Internal Mammary Artery (IMA) in Coronary Artery Bypass Graft (CABG)

These eleven measures represent the component measures for the 0696 composite. In a brief conversation, the Committee noted that the components are necessary to calculate the composite and that the data are already gathered through registry participation. The Committee further noted that if the components were not separately endorsed, each component would need to be reviewed as part of the composite review. The Committee was satisfied there was a justification for the related measures and that burden was minimized.

Public Comment

No public or NQF member comments were provided during the measure evaluation meeting.

Next Steps

NQF will post the draft technical report on March 30, 2020 for public comment for 30 calendar days. The continuous public comment with member support will close on April 20, 2020. NQF will re-convene the Standing Committee for the post-comment web meeting on May 13, 2020.