NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Cost and Efficiency Standing Committee – Fall 2020 Measure Evaluation Web Meeting

The National Quality Forum (NQF) convened the Cost and Efficiency Standing Committee (link to slides) for two web meetings on February 11 and 26, 2021, to evaluate one maintenance measure for endorsement consideration.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF staff reviewed the meeting objectives. The Standing Committee members each introduced themselves and disclosed any conflicts of interest. One Standing Committee member, Danny van Leeuwen, was recused from the measure evaluation discussions and voting because he previously served on a developer-convened Technical Expert Panel (TEP) that informed the development of NQF #2158. Another Standing Committee member, Bijan Borah, MSc, PhD, was recused from voting on the scientific acceptability criteria (i.e., reliability and validity) for NQF #2158 due to his involvement with the NQF Scientific Methods Panel (SMP). Lastly, one Standing Committee member, Dinesh Kalra, was termed *inactive* for this cycle due to a competing engagement.

During both meetings, the quorum required for voting was not achieved (14 out of 20 Standing Committee members). Therefore, the Standing Committee discussed all relevant criteria and voted offline using a web-based voting tool.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and the current NQF portfolio of endorsed measures. There are currently 13 NQF-endorsed measures in the Cost and Efficiency portfolio. Additionally, NQF reviewed the Consensus Development Process (CDP) and the measure evaluation criteria.

A measure is recommended for endorsement by the Standing Committee when the vote margin on all must-pass criteria (Importance, Scientific Acceptability, Use), and overall, is greater than 60 percent of voting members in favor of endorsement. A measure is not recommended for endorsement when the vote margin on any must-pass criterion or overall is less than 40 percent of voting members in favor of endorsement. The Standing Committee has not reached consensus if the vote margin on any must-pass criterion or overall is between 40 and 60 percent, inclusive, in favor of endorsement. When the Standing Committee has not reached consensus, all measures for which consensus was not reached will be released for NQF member and public comment. The Standing Committee will consider the comments and re-vote on those measures during a webinar convened after the commenting period closes.

Measure Evaluation

The Cost and Efficiency Standing Committee evaluated one maintenance measure for endorsement consideration. Pre-evaluation meeting comments from NQF members and the public were also considered by the Standing Committee and can be found in Appendix A. The Standing Committee's deliberations will be compiled and provided in the draft technical report. NQF will post the draft

technical report on April 1, 2021, 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

NQF #2158 Medicare Spending per Beneficiary (MSPB) Hospital (Centers for Medicare & Medicaid Services (CMS)/Acumen, LLC.)

Centers for Medicare & Medicaid Services (CMS)/Acumen, LLC. Representatives at the Meeting

- Sri Nagavarapu, PhD
- David Ruiz, PhD
- Joyce Lam, MPP
- Rose Do, MD

Standing Committee Votes

- Importance to Measure and Report: Total Votes – 16; H-8; M-8; L-0; I-0
- Reliability: Total Votes – 15 (due to SMP member, Bijan Borah, recusal); H-8; M-6; L-1; I-0
    - This measure is deemed as complex and was evaluated by the NQF SMP.
    - The NQF SMP's ratings for Reliability: High (H-7; M-0; L-0; I-0)
    - Since voting was conducted offline using a web-based voting tool, the Standing Committee provided their own vote for reliability rather than be asked to uphold the SMP's rating.
- Validity: Total Votes – 15 (due to SMP member, Bijan Borah, recusal); H-1; M-11; L-3; I-0
    - This measure is deemed as complex and was evaluated by the NQF SMP.
    - The NQF SMP's ratings for Validity: Moderate (H-1; M-6; L-0; I-0)
    - Since voting was conducted offline using a web-based voting tool, the Standing Committee provided their own vote for validity rather than be asked to uphold the SMP's rating.
- Feasibility: Total Votes – 16; H-12; M-4; L-0; I-0
- Use: Total Votes – 16; Pass-16; No Pass-0
- Usability: Total Votes – 16; H-1; M-11; L-4; I-0

Standing Committee Recommendation for Endorsement: Total Votes – 16; Yes-13; No-3

The Standing Committee recommended the measure for continued endorsement. Originally endorsed in 2017, the focus of this measure is to assess the cost to Medicare for Part A and Part B services performed by hospitals and other healthcare providers during an MSPB Hospital episode, which comprises the periods of three days prior to, during, and 30 days following a patient's hospital stay. The Standing Committee questioned the extent to which hospitals could control patient outcomes. They noted that the key improvement opportunities provided by the developer focused on post-acute care settings, which led to questions from the Standing Committee as to whether attribution should be shared between hospitals, nursing homes, and other post-acute settings. They discussed the types and availability of post-acute care services and its driving force in cost variation. The developer responded, stating that the variation within the 30-day period aligns with other Medicare Spending per Beneficiary – Post-Acute Care (MSPB-PAC) measures. The developer explained that providers can exert control over patient outcomes after acute care hospitalization during this period by utilizing their resources and

relationships for referrals. In addition, the risk adjustment allows for comparison of similar episodes within diagnostic categories and exclusion of outlier episodes.

When considering the opportunity for improvement, the Standing Committee commented that the interquartile range was narrow. The developer commented that when looking at the change in improvement from 2017 to 2018, the tail end of the distribution had a difference of 0.09 (from 0.94 in 2017 to 1.03 in 2018). The developer stated that these changes were substantial due to the difference in associated costs. The average episode cost was $22,000, and the 9 percent change equated to almost $2,000 per episode. With around six million episodes, the opportunity for improvement was estimated to be $12 billion. The Standing Committee ultimately agreed that this measure addresses a high-impact/high-resource use area of healthcare.

The Standing Committee noted that this measure was evaluated by the SMP and was given a high rating for reliability and a moderate rating for validity. The Standing Committee did not have any major concerns with respect to reliability and passed the measure on this criterion. Regarding validity, the Standing Committee had concerns about the exclusion of social risk factors from the risk adjustment model. The Standing Committee also acknowledged one public comment that was received for this measure, which raised a similar concern with the lack for social risk factors included in the model. The commenter also raised concern that the current risk adjustment model is not adequate due to the unadjusted and adjusted R-squared results ranging from 0.11 to 0.67. Standing Committee members stated that the difference in performance between hospitals may be masked while adjusting for the within hospital disparity. Standing Committee members also commented that hospitals with more patients affected by social risk factors would have higher costs, which would have adverse effects on their measure performance. The Standing Committee recommended that the developer re-examine how risk factors are entered into their risk adjustment model to include hospital fixed effects, as the 109 risk factors that are included may not be precisely estimated. The developer stated that they attempted an alternative approach to risk adjustment by examining a different model specification for social risk factor testing with dual eligibility included, but this presented challenges that involved the inclusion of another 3,000 estimators of provider effects. The developer further stated that this led to precision error in the estimation prediction. Therefore, the developer did not include social risk factors in the final model. The Standing Committee also questioned how the developer would account for the cost variation from the effects of COVID-19. The Standing Committee noted that race was not a part of risk adjustment and that Black and Brown people have been disproportionately affected by COVID-19. The developer responded, stating that they were working on monitoring 2020 data with CMS. They stated that all claims from January 1, 2020 – June 30, 2020 for a series of measures across certain hospital programs will be excluded. The Standing Committee did not have any further questions or concerns and passed the measure on validity.

The Standing Committee regarded the measure as feasible. In their discussions related to usability and use, the Standing Committee recognized that this measure is currently used in the Hospital Value-Based Purchasing (HVBP) Program. The developer mentioned how they incorporated feedback from users through various avenues, such as annual webinars and help desks. The Standing Committee noted that a survey of users and an empirical analysis of user feedback would be nice to see in the future. Ultimately, the Standing Committee passed the measure on use. Moving to usability, the Standing Committee stated that it was unclear as to what drives variation in public reports for this measure, as the data presented were very high level. The Standing Committee questioned how feasible it would be for a hospital to use measure performance data for improvement and recommended that the reports be more granular. The developer mentioned that reports show patient-level data that are risk-adjusted to allow for comprehension of the case mix. The developer further stated that they compare hospitals'

relationship to median hospital costs to evaluate provider change relative to other hospitals. This comparison would be in consideration of patients' risk level, as it is stratified to account for various conditions. Improvement is measured by achievement in how the measure is used within the program and the change in the hospital's annual performance. The Standing Committee also discussed whether there was a best practice example of a hospital or hospital-based accountable care organization with a performance improvement initiative that showed improvement in cost savings using NQF #2158. One Standing Committee member stated that the Medicare hip/knee replacement bundled program may be the most prominent example of an episode-based program showing cost savings. It was noted that cost savings were achieved by shifting from skilled nursing facilities to home or home health services and by considering costs in conjunction with other outcomes, such as readmission rates, to aid in the evaluation of the quality of care. The Standing Committee did not raise any additional questions or concerns, passed the measure on usability, and recommended this measure for continued endorsement. There was one public comment received that the Standing Committee considered in their evaluation of the measure. This comment focused on the following: (1) how risk adjustment factors, namely social risk factors, were considered, and (2) the rationale for revisions to the measure, namely the change to equally weigh all risk-adjusted hospital episodes by the average ratio of observed to expected costs and the expansion of episodes to include re-hospitalizations within 30 days of discharge of any admission that opens an episode.

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 April 1, 2021, for public comment for 30 calendar days. The continuous public commenting period with member support will close on April 30, 2021. NQF will reconvene the Standing Committee for the post-comment web meeting on June 2, 2021.

Appendix A: Pre-Evaluation Comments

Comments received as of January 28, 2021:

Topic	Commenter	Comment
2158: Medicare Spending Per Beneficiary (MSPB) Hospital	Submitted by American Medical Association	The American Medical Association (AMA) requests that the Standing Committee discuss the revisions made to the measure as described in S.7.2, specifically the change to equally weigh all risk-adjusted hospital episodes by the average ratio of observed to expected costs, and the expansion of episodes to include re-hospitalizations within 30 days of discharge of any admission that opens an episode. No rationale was provided for any of these changes, which makes it difficult for the AMA to provide input and determine whether we agree with the changes. The AMA is particularly concerned that the expansion to include re-hospitalizations will now double count the costs attributed to a hospital. The AMA does not believe that the current risk adjustment model is adequate due to the unadjusted and adjusted R-squared results ranging from 0.11 to 0.67 across the Major Diagnostic Category. The measure is not adequately tested and adjusted for social risk factors. It is unclear why the measure developer would test social risk factors after adjusting for clinical risk factors rather than assessing the impact of both clinical and social risk factors in the model at the same time. These variations in how risk adjustment factors are examined could also impact how each variable (clinical or social) perform in the model and remain unanswered questions. In addition, we note that hospitals measure scores shift when some or all of the social risk factors are applied within the risk model and particularly just over 15% of safety-net hospitals move above or below the delta in Model 13 (Table 2b34b.c Impact of Social Risk Factors). We ask the Standing Committee to carefully consider whether these results impact the ability of the measure to meet the validity criterion. Lastly, we would like to express our appreciation that the measure developer completed correlations with existing hospital quality measures and encourage the measure developer to continue to provide this information for other cost measures.