NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Cost and Efficiency Committee – Post Comment Web Meeting

The National Quality Forum (NQF) convened the Cost and Efficiency Standing Committee for web meetings on October 1, 2020 and October 13, 2020 to review and discuss public comments received for six newly submitted measures. The Committee also revoted on measures where consensus was not reached during the July 10, 2020 measure evaluation meeting. Lastly, the Committee reviewed reconsideration requests for two measures that did not pass on validity during the July 10, 2020 meeting.

During the July 10, 2020 measure evaluation meeting, the Cost and Efficiency Committee evaluated six newly submitted measures.

The Committee recommended one measure for endorsement:

- NQF 3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals (LTACH) (Acumen, LLC)

The Committee did not recommend two measures for endorsement:

- NQF 3561 Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRF) (Acumen, LLC)
- NQF 3574 Medicare Spending Per Beneficiary Clinician (Acumen, LLC)

The Committee did not reach consensus on three measures for endorsement:

- NQF 3563 Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled-Nursing Facilities (SNF) (Acumen, LLC)
    - Consensus Not Reached on Validity
- NQF 3564 Medicare Spending Per Beneficiary – Post Acute Care Measure for Home Health Agencies (HHA) (Abt Associates)
    - Consensus Not Reached on Reliability and Validity
- NQF 3575 Total Per Capita Cost (TPCC) (Acumen, LLC)
    - Consensus Not Reached on Validity

The commenting period for the Cost and Efficiency spring 2020 measure evaluation cycle was August 14, 2020 to September 14, 2020. As of September 14, 2020, NQF received both measure-specific and general draft report comments from NQF members and individuals of the public for Committee review. During this period, the Committee received reconsideration requests submitted by the measure developer for NQF #3561 and NQF #3574.

Day 1: October 1, 2020

During the October 1, 2020 Post-Comment Web Meeting, the Committee did not achieve quorum on the call. Therefore, a recording of the meeting and the offline voting surveys were distributed to members of the Committee.

Welcome, Introductions, and Review of Meeting Objectives

Matthew Pickering, senior director, welcomed the Committee, the developers, and other participants to the web meeting. The Committee co-chairs, Cheryl Damberg and Sunny Jhamnani, provided welcoming

remarks. NQF Sr. Vice President Sheri Winsper also provided welcoming remarks, and Janaki Panchal, manager, conducted a roll call. Matthew Pickering then reviewed the following meeting objectives:

- Discuss & Revote on Consensus Not Reached (CNR) Measures
- Review and Discuss Public Comments
- Discuss & Vote on Reconsideration Requests

For the CNR and reconsideration measures, Matthew Pickering introduced each measure for discussion by presenting a brief overview of the measure, noting any previous concerns that arose during the Committee's initial review on July 10, 2020, and summarized any comments received along with the developer's responses. A summary of the public comments and developer responses are in the comment memo and comment table.

Discussion & Revote on Consensus Not Reached (CNR) Measures

NQF 3563 Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled-Nursing Facilities (SNF) (Acumen, LLC) – *Validity: H-0; M-9; L-6; I-1*

Matthew Pickering introduced the measure and summarized that during the July 10, 2020 proceedings, the Committee raised several concerns regarding the threats to validity including exclusions, alignment of patient risk with SNF payment programs, and the lack of social risk factors within the risk adjustment model. Co-chair Cheryl Damberg opened the floor for the developer and Committee discussion.

The developer stated that the exclusions were appropriate as they were outside of providers' control. The developer stated that certain payment variables were excluded from the risk adjustment model because they are susceptible to potential gaming and manipulation. The developer maintained that the exclusion of social risk factors is not a threat to the measure's validity and further aligned with NQF guidelines for considering risk adjustment factors. Additionally, the developer mentioned that when it tested both the provider- and patient-level variables in the model to examine the marginal effect, the developer noted that the provider-level effect was 17 times larger than the patient-level effect. This indicated to the developer that including only the patient-level variable into the model would mask the disparity between providers based on the proportion of dual eligibility patients. However, Committee members stated that social risk factors should have been included, as the between provider effect would still be evident. The Committee emphasized that risk adjustment does not completely remove or mask all disparities. The Committee also asked NQF's position on the developer's rationale of excluding social risk factors in risk adjustment model. Matthew Pickering explained that NQF provides guidance on how to approach the assessment of social risk factors in the NQF Measure Evaluation Criteria, and the Committee is to determine whether or not to accept the developer's decision based on the evidence and data provided.

Committee members also discussed whether an adequate amount of consideration had been given to the differences in costs within the SNFs and how much of the post-SNF costs could be attributed to care within them. Committee members noted that the current model did not fully capture the various reasons and diagnoses for patient visits, which, in turn, may mask the payment level associated with each patients' reason for their visit. The developer stated that they established clinical categories/themes for each of the settings including post-acute care (PAC), working internally with PAC clinicians to divide all the patients into clinical cohorts prior to calculating the measures in order to address the complexity of the patients.

After the offline votes were received, the Committee did not pass the measure on validity.

NQF 3564 Medicare Spending Per Beneficiary – Post Acute Care Measure for Home Health Agencies (HHAs) (Acumen, LLC) – *Reliability: H-1; M-13; L-2; I-0; Validity: H-0; M-9; L-7; I-0*

Matthew Pickering introduced the measure and summarized that during the July 10, 2020 proceedings, the Committee raised concerns that the reliability statistics for low volume providers were too low for acceptable reliability (0.57), and that several Committee members noted that it may be difficult to differentiate HHAs with smaller number of qualifying episodes. Co-chair Cheryl Damberg opened the floor for the developer and Committee discussion regarding the reliability criterion.

The Committee asked the developer to what extent the reliability for small providers was considered. The developer stated the reliability is considered at different case minimums with varying amounts of reporting entities. The Committee member raised concerns that the reliability cut points were too low. The Committee questioned whether the steward and the developer were predetermining the reliability by choosing the interaction with minimum reporting size and how reliability would be assessed. A Committee member recommended to avoid moving into the space where you're getting more noise than signal in the estimate.

With respect to validity, Matthew Pickering summarized that several Committee members raised concern that the developer reported a low overall risk adjustment R-squared of 0.092. The Committee also raised concerns regarding the developer's exclusion of social risk factors in the overall risk adjustment model, given that these factors were statistically significant, and that the approach to characterizing patient risk for the expected cost is not aligned with the approach to handling payment for HHAs. Co-chair Cheryl Damberg opened the floor for the developer and Committee discussion.

Similar to NQF #3563, the Committee considered the developer's approach to examining the marginal effects of testing both the provider- and patient-level variables in the model. The Committee asked which variables were correlated in the risk adjustment model. The developer explained that the patient-level indicator variable for being dual-eligible and the provider-level proportion of dual-eligible patients were correlated in testing. The developer re-emphasized that the provider-level variable is correlated with the outcome, and it is correlated with the patient-level variable excluded. The developer stated race, ethnicity, and socio-economic status were tested individually and combined. The developer noted that even though they were found to be significant, their overall impact on the measure score was small, showing little change to the R-squared value in the model.

One Committee member recommended the developer include the HHA variable as a random effect variable in the model. The developer responded saying that using random effects would not allow for visibility of the within provider variation, because the random effect is assumed to be uncorrelated with everything else in the model, and therefore, the developer chose to use a fixed effects approach. The developer also illustrated that they attempted to maintain the inclusion of services that are within the scope of the provider and that there are tradeoffs between service exclusions and the R-squared value.

After the offline votes were received, the Committee voted to pass the measure on reliability, but did not pass the measure on validity.

NQF 3575 The Total Per Capita Cost (TPCC) (Acumen, LLC) – *Validity: H-0; M-10; L-5; I-1*

Matthew Pickering introduced the measure and summarized that during the July 10, 2020 proceedings, the Committee raised concern regarding the reported 15.3 percent of episodes being excluded from the measure and concerns with the developer's exclusion of social risk factors in the overall risk adjustment model, given that they were statistically significant. Co-chair Cheryl Damberg opened the floor for the developer and Committee discussion regarding the validity criterion.

The developer informed the Committee that 14.3 percent of the reported 15.3 percent of episode exclusions is due to patients that are not continuously enrolled in Medicare Parts A and B. Therefore, the developer could not observe their cost data. With respect to the risk adjustment of social risk factors, the developer considered adding the Agency for Healthcare Research and Quality Socioeconomic Status (AHRQ SES) Index into the model. However, the developer reported that the AHRQ SES variable had a very small coefficient and did not have an impact on the overall model fit. Therefore, the developer did not include this variable in the model and stated that gender and dual eligibility were included in the model. A question was raised whether race or ethnicity were tested. The developer stated that it did not test for either of these factors.

In addition to empirical validity testing, the developer also encouraged the Committee to consider the face validity testing that was performed on the measure. The developer stated that NQF guidance was utilized for face validity testing. Matthew Pickering reminded the Committee that face validity is an acceptable form of validity testing for new measures and is a minimum requirement, and the highest rating that can be achieved is "moderate" for face validity.

A Committee member requested clarification on whether face validity was done on the results. Matthew Pickering referenced the measure evaluation criteria and stated that that systematic approach for considering the performance scores would include evaluation of results, as performance scores are included. The developer stated that the face validity vote was on the ability of this measure to distinguish systematic performance differences between providers. The developer stated that the measure's summary statistics on how the measure would change in aggregate, including performance scores based on the modifications made to the measure, were provided to their technical expert panel.

After the offline votes were received, the Committee voted to pass the measure on validity.

Review and Discuss the Reconsideration Request

A reconsideration request from the measure developer, Acumen, LLC, was received for NQF #3561 *Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities* on the basis that the measure evaluation criteria were not applied appropriately.

Matthew Pickering explained the reconsideration request process, stating that the Committee would vote on whether they would like to reconsider the measure based on comments received from the public and the rationale provided by the developer. If greater than 60 percentof the Committee votes "yes," the Committee would continue their review of the measure starting with the criterion the measure did not pass. There is no grey zone for reconsiderations. However, since quorum was not achieved during Day 1 of the post comment proceedings, the measure was opened back up for discussion and the Committee proceeded to review NQF #3561, starting with the criterion that the measure did not pass, and voting was captured offline.

NQF 3561 Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities (IRFs) (Acumen, LLC) – *Validity: H-0; M-12; L-5; I-0; Feasibility: H-6; M-13; L-0; I-0; Use: Pass-17; No Pass-2; Usability: H-0; M-16; L-3; I-0; Overall Suitability: Pass-13; No Pass-6*

Matthew Pickering introduced the measure and summarized that during the July 10, 2020 measure evaluation meeting, the Committee raised several concerns regarding the validity of the measure. The Committee has questioned the use of 30-days as the appropriate length of time that IRF can influence downstream care decisions and highlighted the need to empirically evaluate and validate if 30-days post-discharge period is the appropriate length of time to capture complications that can be attributed to IRF care. The Committee also raised concern that the calculation of expected cost is not aligned with

IRF payment programs and how patient risk is accounted for in IRF payment programs. The Committee also raised concerns with the risk adjustment model, specifically the lack of adjustment for social risk factors.

Matthew Pickering stated the developer's rationale for reconsideration was that the evaluation criteria were not correctly applied and that there was inconsistent application of the evaluation criteria that led to a measure not being recommended for endorsement. The developer stated that NQF #3561, NQF #3563, and NQF #3564 purposefully exclude some payment variables, noting that this was an explicit policy decision by the Center for Medicare and Center for Clinical Standards and Quality as excessive spending in these settings has historically been driven by excessive use of therapy and variability in coding of patient status on assessment instruments. The developer mentioned that they did consider such variables and included a wide range of clinical factors, including IRF Rehabilitation Impairment Categories (RICs). However, the developer mentioned that inclusion of such variables would violate NQF guidance, which emphasizes that variables be resistant to gaming, are not indicators of the care provided, and be present at the start of care.

With respect to social risk adjustment, the developer mentioned that adjusting for social risk factors may mask disparities in care, creating a lower standard of care for beneficiaries with higher social risk, and that this could allow for a higher rate of readmissions, complications, etc., among those with high social risk. The developer mentioned that this may be appropriate if such outcomes are outside of a provider's control, but the developer's empirical testing showed that poorer performance for high social risk individuals is closely tied to providers themselves, rather than individual beneficiaries, with especially strong effects in particular settings. The developer argued that due to sign of the relationship between social factors and the measure scores being negative, and inclusion of certain factors may penalize providers for taking on beneficiaries with high social risk. Finally, the developer stated that including these factors in the model had a minimal impact on provider scores. Co-chair Sunny Jhamnani opened the floor for the developer and Committee discussion regarding the validity criterion. The Committee did not have any further questions for the developer.

Matthew Pickering reminded the Committee that since quorum was not achieved, the Committee would need to proceed with the discussion and review of the measure, starting with the feasibility and voting would be captured offline. Co-chair Sunny Jhamnani gave a brief overview of the measure's feasibility, use, and usability. The Committee did not raise any questions or concerns for feasibility, use, and usability.

After the offline votes were received, the Committee voted to pass the measure on validity, feasibility, use, and usability. The Committee also passed the measure on the overall suitability for endorsement.

Day 2: October 13, 2020

During the October 13, 2020 Post-Comment Web Meeting, quorum was achieved on the call. Therefore, Committee voting was held during the call.

Welcome, Introductions, and Review of Meeting Objectives

Matthew Pickering, NQF senior director, welcomed the Committee and participants to the web meeting and provided a brief recap of the October 1, 2020 meeting. Janaki Panchal, NQF manager, the proceeded to conduct a roll call. Matthew Pickering then reviewed the objectives of the meeting:

- Discuss & Vote on Reconsideration Requests for NQF #3574
- Review & Discuss Public Comments for NQF #3562

Discuss & Vote on Reconsideration Request

A reconsideration request from the measure developer, Acumen, LLC, was received for NQF #3574 *Medicare Spending Per Beneficiary Clinician* on the basis that the measure evaluation criteria were not applied appropriately.

Matthew Pickering explained the reconsideration request process, stating that the Committee would vote on whether they would like to reconsider the measure based on comments received from the public and the rationale provided by the developer. If greater than 60 percent of the Committee votes "yes," the Committee would continue their review of the measure starting with the criterion the measure did not pass. If more than 60% of the Committee does not vote "yes," the Committee would not reconsider the measure. There is no grey zone for reconsiderations.

NQF 3574 Medicare Spending Per Beneficiary (MSPB) Clinician – *Reconsideration Vote: Y-5; N-12*

Matthew Pickering introduced the measure and summarized the concerns from the July 10, 2020 measure evaluation discussion. The Committee shared similar concerns with the NQF Scientific Methods Panel regarding the strength of the correlations, noting that the correlation between predicted value and six different clinical themes (e.g., post-acute care (PAC) settings) was low (< 0.10) in all cases except PAC IRF/LTACH, and that the correlation with risk adjusted value and six different clinical themes was also low and negative (-0.18) with PAC Home Health. Lastly, the Committee raised concerns regarding the lack of including social factors within the risk adjustment model. Co-chair Sunny Jhamnani opened the floor for the developer and Committee discussion regarding the validity criterion. The Committee did not have any further questions for the developer.

During the post-comment meeting, Committee members discussed the issues related to the masking of disparities with absence of social risk factors in the model and the low correlation in empirical testing results. The developer stated that the risk adjustment model adjusts for payments variables, like the LTACH measure. The developer stated that any approach that uses provider fixed effects will yield statistically different estimates than approaches that do not eliminate provider variation. The developer stated that their approach shows that providers with large numbers of dual beneficiaries tend to have systematically worse performance. Regarding the correlation analyses, the developer stated that each clinical theme was meant to be evaluated in its clinical context. The developer noted the themes should have lower or negative correlations in cases where they are believed to substitute for more expensive care downstream.

In addition to empirical validity testing, the developer also encouraged the Committee to consider the face validity testing that was performed on this measure. The developer stated that NQF guidance was

utilized for face validity testing and mentioned that the NQF algorithm within NQF's Measure Evaluation Criteria depicts that if face validity is established for a measure, it is an acceptable method to establish validity for new submissions. Matthew Pickering added that within the NQF Measure Evaluation Criteria, while face validity can be an acceptable form of validity for new measures, it is evaluated together with the threats to validity, like that of the risk adjustment model.

The Committee voted not to reconsider measure NQF #3574.

Review & Discuss Public Comments for NQF 3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals (LTACH) (Acumen, LLC)

Matthew Pickering provided a brief overview of the measure and comments received, along with developer's response. A commenter stated that Long-Term Acute Care Hospitals' funding and utilization are controversial. Though the commenter supported the Committee's endorsement of the measure, they believed that high LTACH utilization does not necessarily correlate with higher quality or better outcomes and suspected that there was substantial regional variation. The developer responded that all the post-acute care measures were developed to address the resource use domain of the IMPACT Act which ensures that cost measures will be considered alongside quality measures including assessment-based ones. The developer stated that the LTACH providers involved in the delivery of high-quality care and appropriate discharge planning and post-treatment care coordination would be expected to perform well with these measures, since beneficiaries would likely experience fewer costly related adverse events.

The Committee did not have any dissenting views or opinions based on the developer's response.

NQF Member or Public Comment

Janaki Panchal opened the floor for NQF member and/or public comments. There were no comments provided.

Acknowledge Committee Members with Expiring Terms

NQF staff acknowledged Committee members whose terms end this year: Troy Feisinger, Rachael Howe, Lisa Latts, Jason Lott, Jack Needleman, Janis Orlowski, John Ratliff, and Srinivas Sridhara. NQF staff and the Committee expressed appreciation for the time and knowledge these Committee members have shared over their years of participation on the Committee.

Next Steps

Funmilayo Idaomi, NQF Analyst, presented the next steps, notifying the Committee that NQF will convene the Consensus Standards Approval Committee (CSAC) web meeting on November 17-18, 2020 for review and approval of the measures. Following CSAC review, there will be an appeals period, scheduled from November 23, 2020 through December 22, 2020.