NATIONAL QUALITY FORUM

Memo

October 24, 2019

To: Cost and Efficiency Standing Committee

From: NQF staff

Re: Post-comment web meeting to discuss public comments received and NQF member expression of support

Purpose of the Call

The Cost and Efficiency Standing Committee will meet via web meeting on October 31, 2019 from 2:00 pm – 4:00 pm ET. The purpose of this call is to:

- Review and discuss comments received during the post-evaluation public and member comment period;
- Provide input on proposed responses to the post-evaluation comments;
- Review and discuss NQF members' expression of support of the measures under consideration; and
- Determine whether reconsideration of any measures or other courses of action are warranted.

Standing Committee Actions

1. Review this briefing memo and draft report.
2. Review and consider the full text of all comments received and the proposed responses to the post-evaluation comments.
3. Review the NQF members' expression of support of the submitted measures.
4. Be prepared to provide feedback and input on proposed post-evaluation comment responses.

Conference Call Information

Please use the following information to access the conference call line and webinar:

Speaker dial-in #: 800-768-2983

Access Code: 5593515

Weblink: https://core.callinfo.com/callme/?ap=8007682983&ac=5593515&role=p&mode=ad

Background

During two web meetings both held on June 27, 2019, the Cost and Efficiency Standing Committee evaluated three newly submitted measures. The Standing Committee recommended all three measures for endorsement. The measures recommended for endorsement are:

- 3509 Routine Cataract Removal with Intraocular Lens (IOL) Implantation (CMS/Acumen)
- 3510 Screening/Surveillance Colonoscopy (CMS/Acumen)

- 3512 Knee Arthroplasty (CMS/Acumen)

Comments Received

NQF solicits comments on measures undergoing review in various ways and at various times throughout the evaluation process. First, NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). Second, NQF solicits member and public comments during a 16-week comment period via an online tool on the project webpage.

Pre-evaluation Comments

NQF solicits comments prior to the evaluation of the measures via an online tool on the project webpage. For this evaluation cycle, the pre-evaluation comment period was open from May 8, 2019 and closed on June 19, 2019 for the measures under review. The three comments received were during this period related to validity and risk adjustment.

Post-evaluation Comments

The draft report was posted on the project webpage for public and NQF member comment on August 8, 2019 for 30 calendar days. The Standing Committee's recommendations will be reviewed by the Consensus Standards Approval Committee (CSAC) on November 12, 2019. The CSAC will determine whether or not to uphold the Standing Committee's recommendation for each measure submitted for endorsement consideration. All Committee members are encouraged to attend the CSAC meeting to listen to the discussion. During this commenting period, NQF received three comments from three organizations:

Member Council	# of Member Organizations Who Commented
Consumer	0
Health Plan	0
Health Professional	1
Provider Organization	0
Public/Community Health Agency	2
Purchaser	0
QMRI	0
Supplier/Industry	0

We have included all comments that we received (both pre- and post-evaluation) in the comment table (excel spreadsheet) posted to the Committee SharePoint site. This comment table contains the commenter's name, comment, associated measure, topic (if applicable), and—for the post-evaluation comments—draft responses (including measure steward/developer responses) for the Committee's consideration. Please review this table in advance of the meeting and consider the individual comments received and the proposed responses to each. The Standing Committee's recommendations will be reviewed by the Consensus Standards Approval Committee (CSAC) on November 12, 2019. The CSAC will determine whether or not to uphold the Standing Committee's recommendation for each

measure submitted for endorsement consideration. All Committee members are encouraged to attend the CSAC meeting to listen to the discussion.

Due to the limited number of the comments and the specific nature of the stated concerns, the Committee will discuss each comment and the concerns identified in each. Additionally, measure stewards/developers were asked to respond where appropriate. Where possible, NQF staff has proposed draft responses for the Committee to consider.

Comments and their Disposition

Measure-Specific Comments

3509 Routine Cataract Removal with Intraocular Lens (IOL) Implantation

A commenter raised concerns around the measure's inclusion of drugs with pass-through status because it may serve as a disincentive to surgeons using innovative treatments in cataract surgery. Further, the commenter disagrees with the inclusion of drugs used to treat post-operative conditions after cataract surgery since this treatment is not part of the cataract surgery itself.

Measure Steward/Developer Response:

We appreciate the input from Eyepoint Pharmaceuticals. To provide some context and to recap the discussions from the Standing Committee meeting, we consider services for potential assignment to the measure by gathering expert input on their clinical relatedness to the procedure or condition that the measure focuses on. For drugs on pass-through status, we assess these on a case-by-case basis to ensure that this measure is capturing clinically relevant services. In this case, the expert clinician committee that we convened had considered that an intraoperative drug on pass-through status was appropriate to include as they considered it to be an important source of cost variation. We continue to monitor the inclusion of this intraoperative drug – as we do with other assigned services – as part of ongoing measure maintenance and will make updates to the measure specifications through the maintenance process where appropriate.

Part D drugs, including eye drops, were not included in the cost measure due to the lack of standardized cost data for Part D claims, as the expert committee believed it would not be appropriate to include non-standardized costs in the measure. An additional challenge is that not every Medicare beneficiary is enrolled in Part D. The expert committee discussed the implications of not including Part D data at this point in time and specifically noted that certain intraoperative Part B drugs, which may be substitutes for eye drops, should nonetheless be included in the measure due to the high cost of those drugs. Once Part D standardized costs are available, we will review the inclusion of Part D costs including eye drops.

Proposed Committee Response:

Thank you for your comment. While the Committee understands the commenters' concerns with the inclusion of pass through status and post-operative drugs in the measure, it also believes that episode-based cost measurement should include all relevant costs in the episode, including drugs used during the episode that are clinically

related to the procedure and treatment of the condition during the episode. Therefore, excluding pass through or other drugs used during the episode of care, are appropriate for inclusion based on the definition of the episode and the intent of the measure.

Action Items:

Does the Committee agree with the proposed response?

Does the Committee wish to reconsider its evaluation ratings or endorsement recommendation?

3510 Screening/Surveillance Colonoscopy

Validity Testing. Commenters expressed concerns with various aspects of the validity of the measure. One commenter sought clarity on Committee's response to the Method Panel's concerns with validity testing approach. Concerns with the low r-squared result for the risk adjustment model testing and the developer's decision not to include social risk factors in the model were also raised; other commenters sought clarity from the developers on their specified approach to risk adjustment. Another commenter questioned whether the measure developer demonstrated meaningful differences in performance and highlighted the implications of program design and benchmarking on distinguishing differences in performance on the measure.

Measure Intent. One commenter expressed concerns with the intent of the measure and its alignment with the measure specifications future improvements as it relates to capturing bowel prep issues. The commenter emphasized that the intent of the measure should be to capture costs related to a colonoscopy episode, and not the efficacy of bowel prep.

Measure Steward/Developer Response:

We appreciate ACG's ongoing engagement with this measure, as the input from specialty societies and the clinician community is vital to the creation of valid, clinically sound episode-based cost measures. Our responses below are structured in parallel with the sections in the submitted comment.

Section 2b1 "Validity Testing": Our validity testing is intended to assess whether the measure is accurately measuring what it is designed to measure. This cost measure is intended to measure the costs to Medicare and beneficiaries for a clinician's performance of an episode of care for colonoscopy, including costs under the reasonable influence of the attributed clinician. For this reason, our analyses of empirical validity examine whether clinicians with higher costs in key categories of services also have higher measure scores.

We understand that, beyond the question of the validity of the measure, there are program-level questions about how this measure is related to existing quality measures. A valid cost measure may have a high correlation with a valid quality measure if clinicians who provide high quality care according to the available measures do so at higher cost. A valid cost measure may have a low correlation with a valid quality measure if a clinician who performs procedures with few complications and targeted utilization of services also has high quality outcomes (in terms of complications or other

metrics). In the future, we will investigate such program-level questions to provide additional information to stakeholders.

Section 2b3 "Risk Adjustment/Stratification": As the MIPS Cost Category is currently scored based on percentile rankings of providers, we performed an analysis to look at the degree to which provider percentile rankings are affected by the inclusion of social risk factors. This analysis shows that 90% of TINs would move no more than +-2 percentiles if social risk factors were included, and that the overall Pearson correlation of the measure with and without social risk factors is 0.999.

Risk adjustment for episode-based cost measures should be evaluated in the context of the service assignment rules, which indicate which costs are counted in the measures and which costs are not counted. This means that much of the variation out of the influence of the clinician is captured by service assignment, rather than risk adjustment. This is a point of distinction between this episode-based cost measure and measures that include all costs. A low R-squared does not necessarily indicate that a measure reflects variation unrelated to clinical care, while a high R-squared does not necessarily indicate the opposite; so, the risk adjustment models must be evaluated in concert with the service assignment rules.

The interpretation of R-squared results for the Colonoscopy measure was also discussed during the Standing Committee meeting, where members noted that the R-squared of the model largely reflects how standardized treatment is, and the variability across patients. As such, it is specific to the procedure or condition for which the costs are being evaluated under each measure, and the clinically specific factors that can affect costs within each measure.

Section 2b4 "Identification of Statistically Significant and Meaningful Differences in Performance": An analysis addressing this question can be found in Section 2a2, Reliability Testing. Reliability is a metric of the precision of the measure, or the ability of a measure to distinguish between "low" and "high" performers, where a result close to 1.00 indicates higher reliability, and a result closer to 0.00 indicates lower reliability. Our testing shows that mean reliability for this measure at a 10 case minimum is 0.956 for TINs, and 0.926 for TIN-NPIs. A mean reliability of 0.4 is generally seen as 'moderate', so these reliability results show that the measure has very high precision.

Our validity testing is intended to assess whether the measure is accurately measuring what it is designed to measure. This cost measure is intended to measure the costs to Medicare and beneficiaries for a clinician's performance of an episode of care for colonoscopy, including costs under the reasonable influence of the attributed clinician. For this reason, our analyses of empirical validity examine whether clinicians with higher costs in key categories of services also have higher measure scores.

We understand that, beyond the question of the validity of the measure, there are program-level questions about how this measure is related to existing quality measures. A valid cost measure may have a high correlation with a valid quality measure if clinicians who provide high quality care according to the available measures do so at

higher cost. A valid cost measure may have a low correlation with a valid quality measure if a clinician who performs procedures with few complications and targeted utilization of services also has high quality outcomes (in terms of complications or other metrics). In the future, we will investigate such program-level questions to provide additional information to stakeholders.

We appreciate ASGE's ongoing engagement with this measure, as the input from specialty societies and the clinician community is vital to the creation of valid, clinically sound episode-based cost measures. Our responses below are structured in parallel with the sections in the submitted comment.

IM 1.2: The low minimum scores reflected idiosyncratic values on a few paid claims and the payment standardization algorithm's handling of these values. The payment standardization methodology prices services using Medicare payment rules. In rare situations, a colonoscopy is paid a positive amount but one which is substantially lower than the Medicare Physician Fee Schedule (PFS) amount. In these cases, the payment standardization methodology would use the actual paid amount instead of the Medicare PFS amount. The corrected testing results for the measure show that the actual minimum scores are $461.21 at both TIN and TIN-NPI levels, reflecting the measure as used in MIPS which does not include these extremely low paid colonoscopy cases.

IM 1.4: We will continue to study the effect of accounting for social risk factors, including race/ethnicity, that would allow us to view disparities and provide information to CMS to assist in any potential improvements. We routinely analyze the effect of accounting for social risk factors on risk adjustment models and clinician measure performance for this measure, as well as the other episode-based cost measures we have developed so far. Our analyses to date have shown that accounting for social risk factors in risk adjustment would not meaningfully affect measure performance, and that the correlation of measures with and without inclusion of social risk factors is very high (above 0.99 for both TIN and TIN-NPI level testing).

IM 2.1: The language in the submission form relating to bowel preparation was intended to indicate one potential factor in reducing downstream costs associated with a colonoscopy, and we recognize that other examples of services – including over-utilization of services or complications – may be more compelling examples for this measure. It was not intended to indicate that this measure is designed to capture data specifically around bowel preparation. Thank you for providing us the opportunity to clarify this point.

U 2.3: Thank you for your note, these bullets were created from development materials for the purpose of this submission form. In future, we will edit these for greater clarity."

We appreciate AGA's ongoing engagement with this measure, as the input from specialty societies and the clinician community is vital to the creation of valid, clinically sound episode-based cost measures. We thank the AGA for their support for NQF endorsement of this measure. Please find below responses to the questions in the submitted comment.

Social Risk Factors: As part of the testing process during development and on the final measure, we performed analyses on the impact of social risk factors on the measure. These analyses showed very minimal impacts on measure scores. The Pearson correlation between the measure with and without social risk factors was 0.999 at TIN and TIN-NPI level. We will continue to monitor the potential effect of social risk factors on the measure.

Within- and Between-Clinician Differences: An analysis of the within and between clinician differences in performance is available in Section 2a2, Reliability Testing. Our signal-to-noise analysis in this section looks at the within-clinician and between-clinician performance variance, with a reliability score close to 1.00 indicating higher reliability, and a result closer to 0.00 indicating lower reliability. Our testing shows that mean reliability for this measure at a 10 case minimum is 0.956 for TINs, and 0.926 for TIN-NPIs.

Risk Adjustment Variables: A list of factors used in the risk adjustment model can be found in the publicly available measure specifications at this location: https://qpp-cm-prod-content.s3.amazonaws.com/uploads/344/2019%20Cost%20Measure%20Code%20Lists.zip. In addition, risk factor variables and their respective coefficients and p-values can be found in the testing appendix in the NQF submission packet.

Committee Responses:

To be determined based on Committee discussion.

Action Items:

- Review and discuss the comments and the developers' responses
- Determine committee response to commenters
- Does the Committee wish to reconsider its evaluation ratings or endorsement recommendation for this measure?

NQF Member Expression of Support

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted for endorsement consideration to inform the Committee's recommendations. One NQF member provided an expression of support: See Appendix A.

Appendix A: NQF Member Expression of Support Results

One NQF member provided expressions of support/nonsupport. One of three measures under consideration received support from NQF members. Results for each measure are provided below.

3509 Routine Cataract Removal with Intraocular Lens (IOL) Implantation (CMS/Acumen)

Member Council	Support	Do Not Support	Total
All Councils	0	0	0

3510 Screening/Surveillance Colonoscopy (CMS/Acumen)

Member Council	Support	Do Not Support	Total
Health Professional	1	0	0
All Councils	1	0	0

3512 Knee Arthroplasty (CMS/Acumen)

Member Council	Support	Do Not Support	Total
All Councils	0	0	0