

Measure Feedback Loop

Web Meeting 3 and 4 – Measure Feedback and the NQF CDP Process

Ashlie Wilbon Jean-Luc Tilly Madison Jung Navya Kumar

April 30, 2019 May 7, 2019

Welcome and Introductions

NQF Project Staff

- Ashlie Wilbon, MS, MPH, FNP-C, Senior Director
- Jean-Luc Tilly, Senior Manager, Data Analytics
- Madison Jung, Project Manager
- Navya Kumar, MPH, Project Analyst

Agenda

- Welcome and Roll Call
- Overview of NQF Process and Evaluation Criteria
- Overview of Use and Usability Criteria and Submission Form
- Other Feedback Considered for the Evaluation of Use and Usability
- Challenges and Gaps
- Opportunity for Public Comment
- Next Steps

Meeting Objectives

- Understand NQF's Use and Usability criteria
- Understand and discuss the current feedback channels, target audiences, and inputs to the evaluation of use and usability
- Identify gaps and challenges in current feedback mechanisms
- Develop recommendations for improving the solicitation and collection of feedback for the evaluation of use and usability

Measure Feedback Loop Committee

- Co-chair: Rose Baez, RN, MSN, CPHQ, CPPS
- Co-chair: Edison Machado, MD, MBA
- Constance Anderson, BSN, MBA
- Robert Centor, MD, MACP
- Elvia Chavarria, MPH
- Dan Culica, MD, PhD
- Melody Danko Holsomback
- Anne Deutsch, RN, PhD
- Tricia Elliott, MBA, CPHQ
- Lee Fleisher, MD

- Mark E. Huang, MD
- Joseph Kunisch, PhD, RN-BC, CPHQ
- Claire Noel-Miller, MPA, PhD
- Ekta Punwani, MHA
- Koryn Rubin, MHA
- Elizabeth (Beth) Rubinstein
- Sue Sheridan, MIM, MBA, DHL
- Jill Shuemaker, RN, CPHIMS
- Heather Smith, PT, MPH
- Deborah Struth, MSN, RN, PhD(c)
- Sara Toomey, MD, MPhil, MPH, MSc

Federal Liaisons

CMS

- Maria Durham
- Sophia Chan
- Patrick Wynne
- Melissa Evans

Project Overview

Project Deliverables



Project Overview

Objectives

- **•** To understand outcomes, and what the unintended consequences are, if any
- To understand how a measure actually performs when in use, and what the possible issues or risks are that may be associated with measure implementation
- To help address whether the measure is having its intended effects on improving quality of care and health measurement

Definitions

- Feedback loop
 - » Refers to the process by which feedback from the measure is relayed back to the multistakeholder standing committee members who recommended the measure to be (re-) endorsed or selected for program use.
 - » In previous CDP projects, standing committee members have expressed the need for updates on how a measure has performed after endorsement. This is especially the case for measures that are contentious, and have a chance of impacting certain stakeholders negatively.
- Feedback
 - » Refers to information about measure performance that could be based on quantitative data or qualitative information

Overview of the Consensus Development Process and NQF Evaluation Criteria

Consideration for Endorsement

Consensus Development Process:

Two Cycles Every Contract Year



NQF Measure Evaluation Criteria for Endorsement

NQF endorses measures for accountability applications (public reporting, payment programs, accreditation, etc.) as well as quality improvement.

- Standardized evaluation criteria
- Criteria have evolved over time in response to stakeholder feedback
- The quality measurement enterprise is constantly growing and evolving—greater experience, lessons learned, expanding demands for measures—the criteria evolve to reflect the ongoing needs of stakeholders

5 Major Evaluation Criteria

- Importance to measure and report: Goal is to measure those aspects with greatest potential of driving improvements; if not important, the other criteria are less meaningful (must-pass)
- Reliability and Validity-scientific acceptability of measure properties: Goal is to make valid conclusions about quality; if not reliable and valid, there is risk of improper interpretation (must-pass)
- Feasibility: Goal is to, ideally, cause as little burden as possible; if not feasible, consider alternative approaches
- Usability and Use (must-pass for maintenance measures): Goal is to use for decisions related to accountability and improvement; if not useful, probably do not care if feasible
- Comparison to related or competing measures

Evaluation of Measures Requires Feedback

Importance

Assessment of performance gap, opportunity for improvement

Feasibility

 Assessment of whether there are any significant barriers to implementation

Usability and Use

- Assessment of whether the measure is (or will be) in use in an accountability application
- Assessment of unintended consequences and benefits of the measure
- Assessment of feedback received on the measure and how it was used

Feedback is Considered from Multiple Perspectives and Channels during Evaluation

Public and NQF Membership – CDP Public Comment Period MAP Workgroups and Committee – Decision and rationale from MAP process, Commenting

Developer – CDP Measure Submission Form

> Standing Committee Evaluation of Use and Usability

Public and NQF Membership – NQF Feedback Tool

Evaluation Criteria and the Submission Form



Evaluation Criteria

-Assessed and/or updated on an annual basis

-Updates reflect evolving measurement science, measurement landscape, and stakeholder feedback

-Reflects best practices in measure development

-Updates must be reviewed and approved by Consensus Standards Approval Committee (CSAC)

Submission Form

-Balances the need to collect necessary information and developer burden

-Only collects information needed to evaluate criteria

-Updates managed by NQF Staff



Overview of Use and Usability Criteria

Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policymakers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

History of the Usability and Use Criteria in the NQF CDP Process

Timeline and changes to the Use and Usability Criteria over time

- Usability: Extent to which intended audiences (e.g., consumers, purchasers, providers, policymakers) can understand the results of the measure and find them useful for decision-making. (2011)
- Usability Report (2012)
- "Usability" split into "Use and Usability" (2013)
- Changed Use to must-pass for maintenance measures (2017)

Principles from Prior Usability Report

- Performance measurement facilitates achieving the goal of high-quality, efficient healthcare for all Americans through two pathways: 1) changes in care initiated by healthcare providers and 2) accountability/selection by making information available to consumers, referring clinicians, and others involved in selecting clinicians and providers.
- NQF strives to endorse measures that are useful for *both* accountability and improvement to maximize their influence on progress toward the goal of high-quality, efficient healthcare for all Americans.
- To achieve maximal effect on quality healthcare and health, over time, NQF-endorsed measures should be used in all applications for which they provide useful information.
- Public disclosure of performance results not only is necessary for some types of selection such as consumer choice, but also ensures accountability and provides external motivation for performance improvement.

Principles from Prior Usability Report

- Measure developers may not be responsible for implementing performance measures for accountability/selection or quality improvement programs and may not have access to the required data or information about measure use.
- The NQF criteria of Importance to Measure and Report and Scientific Acceptability of Measure Properties ensure that a measure is potentially useful for a variety of applications.
- Measures can be more or less useful to intended audiences depending on the conditions of implementation for a specific purpose (e.g., if reporting methods or classification methods obscure differences in performance).
- The NQF criterion of Feasibility, particularly regarding the data required to implement a performance measure, also influences usability. However, feasibility issues may be mitigated or the benefit of measuring performance may outweigh associated burden.

Recommendations from Prior Usability Report

- Evaluate *potential* usability for new measures and *actual* use and performance results of measures subject to endorsement maintenance.
- Set expectations for timeframes to achieve use in accountability applications and public reporting, but allow flexibility.
- The benefits of measurement in terms of facilitating improvement should outweigh evidence of unintended negative consequences.
- Address Usability and Use last in the hierarchy of the four major criteria because if the other criteria are met, then a measure should be usable.
- Usability and Use is not a must-pass criterion.

No longer true for maintenance measures

Discussion

- Do these principles still hold true?
- Are there any additional principles the Committee believes would be helpful in guiding this work and future recommendations?
- Should the recommendations be reconsidered given the evolving measurement landscape?

Use and Usability Criteria Domains

Use

- Accountability and Transparency
- Feedback from those being measured and others

Usability

- Improvement
- Unexpected Findings (unintended consequences and benefits)

Measure Submission Form

- All information submitted by the measure developer/steward
- 598 measure submissions since December 2016
- Themes:
 - Blank or N/A
 - Summary of process
 - No feedback received
 - Feedback received are clarification questions for specifications, ICD codes, methodology, suggestions
 - Unintended consequences
 - Implementation burden
 - Support measure/agree with need for measure
 - Clarification questions

Use and Usability

- Use* Current and planned use of measure
 - Accountability and Transparency: Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement

*Must pass for maintenance measures

Key Questions for Evaluating Use pg 29 of 2018 NQF Measure Evaluation Criteria and Guidance

Accountability/Transparency

- Is it an initial submission with a credible plan for implementation in an accountability application?
- Is the measure used in at least one accountability application (within three years for maintenance measures)?
- Are the performance results publicly reported within six years (or the data on performance results are available)?
- If any of the above answers are "No":
 - What are the reasons (e.g., developer/steward, external factors)?
 - Is there a credible plan for implementation and public reporting?

Use: Submission Form

Accountability and Transparency

 4.1. Current and Planned Use (check all the current and planned uses; for any current uses that are checked, provide a program name and URL for the specific program)

Intended Use	Specific Plan for Use	Current Use	For current use, provide Program Name and URL
a. Payment Program	0	0	
b. Professional Certification or Recognition Program	0	0	
c. Public Health/Disease Surveillance	0	0	
d. Public Reporting	0	0	
e. Quality Improvement (external benchmarking to organizations)	0	0	
f. Quality Improvement (Internal to the specific organization)	0	0	
g. Regulatory and Accreditation Programs	0	0	

h. Not in use	0
i. Use Unknown	0

Use: Submission form

Accountability and Transparency

- 4a1.1. For each CURRENT use, checked above provide:
 - Name of program
 - Program sponsor
 - Purpose
 - Geographic area
 - Number and percentage of accountable entities and patients included
 - Level of measurement
 - Setting
- 4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?
- 4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes—any accountability application within 3 years and publicly reported within 6 years of initial endorsement.

Current State

- Since many developers do not implement their measures, it is possible that there are programs in which the measures are used without the developers' knowledge.
- At least a few developers have decided not to submit for re-endorsement due to a must-pass criterion
- Few measures have failed to obtain endorsement due to this requirement
- Some measures recommended as suitable for endorsement even without passing this criterion
 - Justification and/or plan accepted, even though this is not technically an option

Current State: Use Data



N=513 measure submissions

Discussion

- What expectations should there be for developers to submit use information via the submission form?
- Are there other sources or stakeholders that should be solicited for use information? By the developer? By NQF?
- Does it still make sense for Use to be must-pass for maintenance measures?
 - If so, should timelines be modified?
 - » Used in at least one accountability application within three years after initial endorsement
 - » Publicly reported within six years after initial endorsement
- Should there be an option for justification and/or plan?
 - If so, would this apply to use in accountability programs? To public reporting?

Use and Usability

Use*

Feedback by those being measured or others: Those being measured have been given results and assistance in interpreting results; those being measured and others have been given opportunity for feedback; the feedback has been considered by developers.

*Must pass for maintenance measures

Key Questions for Evaluating Use pg 29 of 2018 NQF Measure Evaluation Criteria and Guidance

Feedback on the measure by those being measured or others

- Summary of the Feedback?
- How was it considered and adjudicated?
- How was feedback obtained/provided?
- Who provided feedback?

Use: Submission Form

Feedback on the measure by those being measured or by others

- 4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.
 - How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.
- 4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.
- 4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others
 - Describe how feedback was obtained.
- 4a2.2.2. Summarize the feedback obtained from those being measured.
- 4a2.2.3. Summarize the feedback obtained from other users.
- 4a2.3. Describe how the feedback has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not
Examples

- "Questions received through the Policy Clarification Support system have generally centered around clarification on the interval between HPV vaccine doses and allowable documentation as proof of vaccination."
- "During a recent public comment session, a majority of comments from measured entities supported updates to the measure to align with the latest clinical recommendations."
- "There have been no significant issues related to the clarity or feasibility of implementing the measure specifications."

Examples

- "Summary of Questions or Comments from Hospitals submitted through the Q & A process:
- For the HF EDAC measure, we have received the following inquiries from hospitals since the completion of measure maintenance in December 2016:
 - 1. Requests for detailed measure specifications including Condition CategoCC-to-ICD-9 code crosswalks, and ICD-9 and ICD-10 codes used to define the measure cohort or in the risk-adjustment model;
 - 2. Requests for the SAS code used to calculate measure results;
 - 3. Queries about how cohorts and outcomes are defined;
 - 4. Queries about how to calculate the measure and to interpret the statistical model including the interpretation of coefficients for risk variables; and
 - 5. Requests for hospital-specific measure information, such as data included in the HSRs; and,
 - 6. Queries about the implementation of EDAC measures in CMS' programs."

Current State

- The summary of the feedback is often blank/not applicable or describes the process in which feedback is collected but not the results
- Often developers do not have the ability to collect feedback because they are not the measure implementers

Discussion

- What expectations should there be for developers to submit complete and meaningful data in response to the questions on the submission form?
- What guidance might there be for developers in collecting and submitting more meaningful data on feedback received for consideration for NQF endorsement?
- Should there be an option for justification and/or plan to solicit and collect feedback on their measure?

Use and Usability

Usability

 Improvement: Progress toward achieving the goal of highquality, efficient healthcare for individuals or populations is demonstrated.

Key Questions for Evaluating Usability pg 30 of 2018 NQF Measure Evaluation Criteria and Guidance

Improvement

- Is it an initial submission with a credible rationale for improvement?
- Has improvement been demonstrated (performance trends, numbers of people receiving high-quality, efficient healthcare)?
- If any of the above answers are "No":
 - What are the reasons?
 - Is there a credible rationale describing how the performance results could be used to further the goal of facilitating highquality, efficient healthcare for individuals or populations?
 - Is the measure used in quality improvement programs?

Usability: Submission Form

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.²² If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

Discussion

- How should developers approach gathering this information for measure submitted for consideration?
- What other channels, organizations, stakeholders should be targeted for soliciting this information?

Usability and Use

Usability

Benefits outweigh the harms: The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations

Key Questions for Evaluating Usability pg 30 of 2018 NQF Measure Evaluation Criteria and Guidance

Benefits Outweigh the Harms

- What were the (potential) negative or unintended consequences?
- What were the (unanticipated) benefits?
- Do the unintended negative consequences outweigh the benefits?

Usability: Submission Form

Benefits Outweigh the Harms

- 4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.
- 4b2.2. Please explain any unexpected benefits from implementation of this measure.



Is there other information about unintended consequences that should be included this evaluation?

- Who would be the target audience for soliciting this information?
- How could it be collected and integrated into the endorsement considerations?

Feedback Considered in the Evaluation of Other Criteria

Importance: Opportunity for Improvement

- 1b.2. Provide performance scores on the measure as specified
 - » This information is also used to address the subcriterion on improvement (4b1) under Usability and Use.
- Feasibility: Data Collection Strategy
 - 3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.
 - IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Discussion

- Are there other domains of feedback and information that should be collected and considered in the evaluation of measures for endorsement?
 - Current evaluation domains include:
 - » Accountability and transparency
 - » Improvement
 - » Benefits outweigh the harms
 - » Feedback from those being measured and others
 - » Data collection strategy
 - » Opportunity for Improvement



The information submitted for evaluation of the Use and Usability criteria is often sparse which has led to Committees "glossing over" its evaluation:

- How might we structure the evaluation of use and usability or the criteria for more meaningful discussion and evaluation?
- Currently Usability is not must pass; should this be reconsidered?

Other Feedback Considered for Evaluation of Use and Usability

Feedback is Considered from Multiple Perspectives and Channels during Evaluation

Public and NQF Membership – CDP Public Comment Period MAP Workgroups and Committee – Decision and rationale from MAP process, Commenting

Developer – CDP Measure Submission Form

> Standing Committee Evaluation of Use and Usability

Public and NQF Membership – NQF Feedback Tool

Summary of Other Feedback Received through NQF processes

Tool/Process	Stakeholder Submitting Feedback	Number	Collection Period
NQF Measure Feedback tool	Public and NQF members	19 comments	2016-Present
CDP Public Comment	Public and NQF members	228 comments	September 2017- Present
MAP Recommendations, Comments	MAP Workgroups and Committee	148 measures reviewed	2016-2019

NQF Measure Feedback Tool

- Public and NQF members can submit and view comments via the NQF homepage and QPS website
- Open text box
 - "We are committed to actively seeking implementation experiences on measures from all stakeholders. Please provide any unexpected findings (positive or negative) during implementation of these measures including unintended consequences or unintended benefits."
- 19 comments received since 2017

Themes:

- Support for measure
- Request for off-cycle review of a measure
- Feedback on specifications
- Concerns about implementation
- Clarification questions

Feedback Tool: Examples

- "This measure is important for the care, safety and wellness of those individuals in the target population served by this measure."
- "It's an excellent measure. However, it is limited by losing a significant amount of relevant data because it excludes multiple ASAM residential treatment levels of care. Many states and other entities would benefit greatly with more accurate data if that observation was considered by the reviewing committee to include residential levels of care in the next update of that measure."

Discussion

- This tool is significantly underutilized:
 - What strategies should NQF employ to increase use of the feedback tool?
 - » Online forum?
 - » Communication strategies?
 - Who should be the target audience for outreach to improve use of this tool?
- Is a free text field sufficient? Should feedback be collected in other formats?

Public Commenting in the CDP

- NQF solicits comments on how NQF-endorsed measures are being used in the field to inform the committee for evaluation of the Usability and Use criterion.
- NQF solicits comments from both NQF members and the public during each evaluation cycle via one continuous 16-week public commenting period.
- All comments received through these mechanisms at least one week prior to the committee evaluation meeting are provided to the standing committee for consideration during the meeting.
- NQF ensures the measure steward/developer receives the submitted comments in a timely manner to prepare for the committee evaluation meeting.
- Measure stewards/developers are not required to provide written responses to the pre-evaluation comments received prior to the measure evaluation meeting.

Public Commenting in the CDP

- 228 comments received since 2016
- Themes:
 - Unintended consequences
 - Implementation burden
 - Issues/suggestions with measure specifications
 - Support measure/agree with need for measure
- Solicited in an open text box field
- Commenting has declined over the last 2-3 years as evaluation activity has increased

Public Commenting: Examples

- "This is an important overall assessment of quality that can assist providers, payors and patients in evaluating the performance related to CABG procedures."
- "FAH is concerned with the limited direct evidence provided to demonstrate that care planning can drive improvements in patient outcomes. FAH also agrees with the Standing Committee's determination that the measure did not meet the scientific acceptability criteria, particularly given the low agreements rates found during reliability testing and the lack of support that the measure's performance score represents true performance as demonstrated in the face validity results. FAH agrees that this measure does not meeting the measure evaluation criteria"

Discussion Questions

- What strategies should be considered for increasing public and member engagement in NQF public commenting?
- Who should be targeted for potential outreach efforts?
- Should the collection of feedback be more prescriptive (rather than an open text box)?

The Measure Applications Partnership (MAP)

- Inform the selection of performance measures to achieve the goal of improvement, transparency, and value for all
- Provide input to HHS during pre-rulemaking on the selection of performance measures for use in public reporting, performance-based payment, and other federal programs
- Identify gaps for measure development, testing, and endorsement
- Encourage measurement alignment across public and private programs, settings, levels of analysis, and populations to:
 - Promote coordination of care delivery
 - Reduce data collection burden

MAP Process Feedback

- Information from the MAP process is included in the measure summary for standing committee consideration
 - Public comments
 - Summary of MAP discussion
 - Based on MAP Workgroup and Committee decisions and rationales for Measures Under Consideration
- 148 measures since 2016-2017 MAP cycle

MAP Process

Themes

- Identifying potential unintended consequences
- Conditions for recommendations
 - NQF endorsement
 - Adjustments to specifications
 - Recommendations for additional testing
- Support of need for the CMS program
- Reasons for not supporting the measure for the CMS program or recommending revisions

MAP Feedback: Examples

- "MAP did not support this measure for rulemaking with the potential for mitigation, which would include specifying and testing the measure at the clinician level of analysis. Additionally, MAP noted the revised measure and testing should be submitted for NQF endorsement review. MAP highlighted the clinical importance of this measure. However, MAP cautioned there is a need for a review with more detailed specifications while considering variability of benefits (i.e., reimbursement for vaccinations), vaccine shortages, data availability/feasibility, and more clarity into the timeframe of reporting, and MAP noted that the composite measure required internal harmonization of its component parts."
- "MAP conditionally supported this measure pending NQF endorsement. MAP members emphasized the importance of multimodal pain management strategies in the light of the current opioid epidemic."

Discussion

- Are there other types of information that should be systematically collected from MAP proceedings to inform endorsement decisions?
- Should other sources for feedback be considered for the evaluation of use and usability of the measures (e.g., CMS impact reports, Jira) as part of MAP deliberations or CDP?

Challenges and Gaps

Feedback is Considered from Multiple Perspectives and Channels during Evaluation

Public and NQF Membership – CDP Public Comment Period MAP Workgroups and Committee – Decision and rationale from MAP process, Commenting

Developer – CDP Measure Submission Form

> Standing Committee Evaluation of Use and Usability

Public and NQF Membership – NQF Feedback Tool

Discussion

- Are there other channels for soliciting or collecting feedback that should be considered for the evaluation of use and usability?
- How might the information collected in the CDP process be shared with other stakeholders in the quality measurement enterprise?

Opportunity for Public Comment

NATIONAL QUALITY FORUM

Next Steps

 CDP Use and Usability Document Draft Report
 Post for 14-day public and NQF member comment period from May 29 to June 11, 2019

CDP Use and Usability Document Final Report
 Due to CMS on July 3, 2019

 Web Meeting 5 – Options for Piloting the Measure Feedback Loop

July 24, 2019, 1-3 pm ET

Overview of Meeting Timeline

Meeting	Date
Web Meeting 3 and 4: Measure Feedback and the NQF CDP Process, Part 1 and 2 [3 hours each]	April 30, 2019, 2-5 pm ET May 7, 2019, 2-5 pm ET
Web Meeting 5: Options for Piloting the Measure Feedback Loop, Part 1 [2 hours]	July 24, 2019, 1-3 pm ET
Web Meeting 6 and 7: Options for Piloting the Measure Feedback Loop, Parts 2 and 3 [2 hours each]	September 3, 2019, 2-4 pm ET September 5, 2019, 2-4 pm ET
Web Meeting 8: Implementation Plan [2 hours]	November 19, 2019, 2-4 pm ET
Web Meeting 9: Project Wrap-Up [2 hours]	January 16, 2020, 1-3 pm ET

Project Contact Information

- Email: <u>measurefeedback@qualityforum.org</u>
- NQF phone: 202-783-1300
- Project page: <u>https://www.qualityforum.org/Measure_Feedback_Loop</u> <u>.aspx</u>
- SharePoint: <u>http://share.qualityforum.org/Projects/MeasureFeedbac</u> <u>kLoop/SitePages/Home.aspx</u>

