

MEASURE APPLICATIONS PARTNERSHIP

Process and Approach for MAP Pre-Rulemaking Deliberations, 2016-2017

FINAL REPORT
FEBRUARY 2017



NATIONAL
QUALITY FORUM

This report is funded by the Department of Health and Human Services under contract HHSM-500-2012-00009I, Task Order HHSM-500-T0018.

CONTENTS

SUMMARY	2
<hr/>	
PROCESS AND APPROACH	2
Overall Approach	2
NQF Member and Public Comment Periods	3
<hr/>	
IMPROVEMENTS THIS YEAR	4
Refinements to MAP Decision Categories	5
Refinements to Preliminary Analysis	6
Improvements to MAP/CDP Integration	8
Review of Current Program Measure Sets	8
Feedback Loop Pilot	9
<hr/>	
BACKGROUND ON RECOMMENDATIONS	10
Voting Process	10

SUMMARY

- In the 2016-2017 pre-rulemaking cycle, MAP examined 71 unique measures for potential use in 16 different federal health programs.
- Process improvements in this cycle included: (1) addition of the *refine and resubmit* decision category for all measures under consideration, (2) refinements to the NQF staff preliminary analysis of measures, (3) removal of the under development voting pathway, and (4) review of current program measure sets. This year, MAP also pilot tested a feedback loop with the PAC/LTC workgroup.

During the pre-rulemaking review cycle, the federal government looks to the Measure Applications Partnership (MAP), a public-private partnership convened by NQF, to advise on the selection of measures for Centers for Medicare & Medicaid Services (CMS) quality initiative programs. Under statute, the Department of Health and Human Services (HHS) is required to publish annually a list of measures under consideration (MUCs) for future federal rulemaking and to consider MAP's recommendations about the measures during the rulemaking process. The annual pre-rulemaking

process affords MAP the opportunity to review the measures under consideration for federal rulemaking and provide upstream input to HHS in a strategic manner. Over the course of the review process, MAP incorporates measure use and performance information into MAP decision making and provides specific recommendations about the best use of available measures and filling measure gaps. This year, MAP also conducted a holistic review of the current program measure sets to provide recommendations on measures for removal.

PROCESS AND APPROACH

Overall Approach

During the 2016-2017 pre-rulemaking period, MAP used a four-step process to analyze and select measures.

1. **Provide program overview.** Using CMS critical program objectives and the NQF Measure Selection Criteria, NQF staff developed a framework for each program measure set in order to organize each program's finalized measure set.
2. **Review current measures.** MAP used the program measure set frameworks to better understand the current measures in the program and identify gaps and other areas of need. MAP reviewed the current measures to help determine how well the measures under consideration might fit into the program.
3. **Evaluate measures under consideration.** MAP used the Measure Selection Criteria and a defined decision algorithm to determine whether the measures under consideration would enhance the program measure sets. Staff performed a preliminary analysis on each measure under consideration using the MAP-approved **preliminary analysis algorithm**. The MAP workgroups discussed their recommendations for each measure under consideration during December in-person meetings. The MAP Coordinating Committee

finalized the recommendations during their January in-person meeting.

4. **Provide feedback on current program measure sets.** MAP reviewed the current measure sets to offer input on how to strengthen them, address gaps, and make recommendations for future removal of measures.

Review Needs and Objectives for Federal Health Programs Under Consideration

In October, MAP workgroups convened via web meeting to consider each program in its setting with the goal of identifying its specific measurement needs and critical program objectives. At an October web meeting, the PAC/LTC workgroup pilot tested the MAP feedback loop. The feedback loop aims to address questions based on stakeholder concerns: Has a measure been submitted for NQF endorsement? What were the results of review by the Endorsement and Maintenance Standing Committee? Is a measure performing as expected? Has a developer updated a measure to address MAP conditions of support? This feedback aims not to change MAP's prior recommendations about a measure, but to give CMS and NQF a chance to update MAP on the development and endorsement of selected measures.

Review Measures Under Consideration and Program Measure Sets

MAP workgroups met in person in December to evaluate the measures under consideration and made recommendations about their potential use in federal programs. MAP reviewed 71 unique measures for potential inclusion in 16 federal health programs. Since some measures were considered for multiple programs, MAP made 74 recommendations on measures within in a particular program.^a For 2016-2017, there were no measures under consideration for two hospital programs. Nevertheless, MAP discussed the current measure sets for these

programs to provide guidance on how to strengthen the measure sets. To assist in their deliberations, MAP members received detailed materials, encompassing all measures and their specifications, preliminary analysis of the measures, and any public comments received. In January, the MAP Coordinating Committee reviewed the workgroup recommendations as well as the public and member comments received on those recommendations. Following deliberations, the Coordinating Committee finalized MAP's recommendations for consideration by HHS.

NQF Member and Public Comment Periods

To encourage early input, MAP has a formal process in which stakeholders can provide feedback on individual measures immediately after HHS provides the year's list of measures under consideration. MAP workgroups took these public comments into account when reviewing the measures under consideration in December. NQF provided stakeholders with a second opportunity for public comment following the preliminary recommendations from the MAP workgroups. Stakeholders provided feedback on the individual workgroup decisions and broader measurement guidance for federal programs. The MAP Coordinating Committee considered these comments when approving the final recommendations on measures to be supported for rulemaking in the federal programs. Throughout the MAP workgroup meetings in December and the MAP Coordinating Committee meeting in January, NQF members and any interested parties from the general public had several opportunities to comment on the list of measures under consideration, on individual workgroup decisions, and on broader measurement guidance for federal programs prior to MAP discussion and voting on the measures under consideration. In prior years, comments were generally made after decisions had already been made. As part of NQF's commitment to transparency, all comments are posted on the NQF website for public reference.

^a The official Measures Under Consideration (MUC) list published on November 21, 2016, contained 97 unique measures under consideration for 16 federal health programs. Three measures were considered for more than one program. CMS withdrew 26 measures from consideration.

IMPROVEMENTS THIS YEAR

Reviewing measures early in their lifecycle has provided a major challenge for MAP. The majority of the measures under consideration in a given year have not been reviewed for NQF endorsement, and some measures under consideration may still be in development and testing. MAP members have asked for an opportunity to better understand what happens to a measure under consideration after MAP reviews it. MAP members want to know: (1) What does the current program measure set include? (2) How is measure development progressing? (3) What are the results of endorsement reviews of measures

included in the programs for which MAP provides input? Tables 1 and 2 below demonstrate the extent of the challenge in selected programs for which MAP reviews measures.

To address this challenge, NQF used feedback from external stakeholders, MAP members, and NQF members to build upon the process improvements from the 2015-2016 pre-rulemaking cycle. This section summarizes several major improvements resulting from that effort to restructure this work, improve the process for those involved in deliberations, and strengthen the deliverables.

TABLE 1. MEASURES REVIEWED BY MAP THAT ARE NOT FULLY DEVELOPED

Program	2014: Not fully developed	%	2015: Not fully developed	%
Hospital IQR	13 of 28	46%	4 of 15	27%
Hospital OQR	6 of 17	35%	0 of 2	0%
LTC Hospital Quality Reporting	-		7 of 7	100%
SNF QRP	-		11 of 11	100%
MSSP	64 of 116	55%	0 of 5	0%
PQRS/MIPS	59 of 96	61%	57 of 61	93%

TABLE 2. NQF-ENDORSED MEASURES ON THE 2015 MUC LIST FOR SELECTED PROGRAMS

Program	Measures on MUC list 2015	
	# NQF-endorsed	# submitted to NQF for endorsement
Hospital IQR	4 of 15 endorsed	2 of 15 submitted
Hospital OQR	0 of 2 endorsed	1 of 2 submitted
LTC Hospital QR	0 of 7 endorsed	0 of 7 submitted
SNF QRP	2 of 11 endorsed	1 of 11 submitted
MSSP	2 of 5 endorsed	1 of 5 submitted
MIPS	2 of 61 endorsed	1 of 61 submitted

Refinements to MAP Decision Categories

MAP reaches a decision about every measure under consideration. The decisions are standardized for consistency. For the 2016-2017 pre-rulemaking cycle, MAP refined the standard decision categories based on feedback from MAP members and the public. Specifically, MAP eliminated the separate pathway for measures

under development that was used in previous pre-rulemaking cycles. However, MAP received feedback that this may set a lower standard for measures early in development and that MAP does not have another opportunity to provide feedback on the measures before they may be implemented.

To address these concerns, MAP now reviews all measures using the same decision categories. Table 3 outlines the decision categories and the evaluation criteria used for each category.

TABLE 3. MAP DECISION CATEGORIES AND EVALUATION CRITERIA

Decision Category	Evaluation Criteria
Support for Rulemaking	The measure is fully developed and tested in the setting where it will be applied and meets assessments 1-6 of the MAP Preliminary Analysis Algorithm. If the measure is in current use, it also meets assessment 7.
Conditional Support for Rulemaking	The measure is fully developed and tested and meets assessments 1-6. MAP will provide a rationale that outlines the conditions (e.g., NQF endorsement) based on assessments 4-7 (reference Table 2 below) that should be met. However, CMS should meet the MAP-specified conditions prior to proposing for rulemaking. CMS may address the MAP-specified conditions without resubmitting the measure to MAP prior to rulemaking.
Refine and Resubmit Prior to Rulemaking	<p>The measure meets assessments 1-3, but needs modifications. A designation of this decision category assumes at least one assessment 4-7 is not met. MAP will provide a rationale that outlines each suggested refinement (e.g., measure is not fully developed and tested OR there are opportunities for improvement under evaluation).</p> <p>Ideally the modifications suggested by MAP would be made before the measure is proposed for use. However, the Secretary retains policy discretion to propose the measure. CMS may address the MAP-specified refinements without resubmitting the measure to the MAP prior to rulemaking. CMS may nonformally, without deliberations and voting, review these refinements via the “feedback loop” with the MAP. These updates may occur during the web meetings of the MAP workgroups scheduled annually in the fall.</p>
Do Not Support for Rulemaking	The measure under consideration does not meet one or more of assessments 1-3.

The evaluation criteria for the decision categories are designed to align with the assessments of MAP's preliminary analysis.

The Refine and Resubmit Prior to Rulemaking category was added for the 2016-2017 pre-rulemaking cycle. MAP added this category to preserve its ability to express support for the concept of a measure under consideration while noting that significant changes may be needed prior to its implementation. The category differs from the Conditional Support for Rulemaking category by signaling that a larger change is needed to the measure under consideration or that the measure under consideration has not completed development and testing.

MAP is committed to the scientific integrity of the measures used in accountability programs but has historically had limited information about the reliability and validity of the measures under consideration. MAP created this decision category to signal a desire for stronger feedback loops between CMS, MAP, and measure developers as well as greater integration with the NQF Consensus Development Process (CDP). MAP members have expressed a desire to understand more about what happens to a measure under consideration after MAP's review, particularly when MAP recommends potential improvements to the measure or the measure has not yet completed testing.

MAP recognizes its advisory role to HHS and that the Secretary retains the ability to propose or not propose any measure under consideration.

However, through the addition of the Refine and Resubmit Prior to Rulemaking category, MAP hopes to establish a pathway to receive feedback on how its recommendations have been addressed. NQF has developed a three-part plan to achieve this goal: (1) the pilot of a feedback loop between MAP and CMS, (2) better integration of the endorsement and selection processes, and (3) review of the current program measure sets. Additional details on these improvements can be found below.

Refinements to Preliminary Analysis

To support members for decisions on individual measures, staff provided a preliminary analysis of all measures under consideration based on a pre-defined and standard algorithm derived from the MAP Measure Selection Criteria and other prior guidance. The preliminary analysis is based on the identified critical program objectives and is intended to provide MAP members with a succinct profile of each measure and to serve as a starting point for MAP discussions. The preliminary analysis algorithm asks a series of questions about each measure under consideration (MUC). This algorithm was revised for the 2016-2017 pre-rulemaking cycle to clarify the information to provide to MAP members and to allow for better integration with the NQF CDP. Table 4 outlines the assessments of MAP's preliminary analysis algorithm.

TABLE 4. MAP PRELIMINARY ANALYSIS ALGORITHM

Assessment	Definition	Outcome
<p>1. The measure addresses a critical quality objective not adequately addressed by the measures in the program set.</p>	<ul style="list-style-type: none"> • The measure addresses the broad aims and one or more of the six National Quality Strategy priorities; or • The measure is responsive to specific program goals and statutory or regulatory requirements; or • The measure can distinguish differences in quality, is meaningful to patients/consumers and providers, and/or addresses a high-impact area or health condition. 	<p>Yes: Review can continue.</p> <p>No: Measure will receive a Do Not Support.</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>
<p>2. The measure is evidence-based and is either strongly linked to outcomes or an outcome measure.</p>	<ul style="list-style-type: none"> • For process and structural measures: The measure has a strong scientific evidence base to demonstrate that when implemented it can lead to the desired outcome(s). • For outcome measures: The measure has a scientific evidence base and a rationale for how the outcome is influenced by healthcare processes or structures. 	<p>Yes: Review can continue</p> <p>No: Measure will receive a Do Not Support</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>
<p>3. The measure addresses a quality challenge.</p>	<ul style="list-style-type: none"> • The measure addresses a topic with a performance gap or addresses a serious reportable event (i.e., a safety event that should never happen); or • The measure addresses unwarranted or significant variation in care that is evidence of a quality challenge. 	<p>Yes: Review can continue</p> <p>No: Measure will receive a Do Not Support.</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>
<p>4. The measure contributes to efficient use of measurement resources and/or supports alignment of measurement across programs.</p>	<ul style="list-style-type: none"> • The measure is either not duplicative of an existing measure or measure under consideration in the program or is superior to an existing measure in the program; or • The measure captures a broad population; or • The measure contributes to alignment between measures in a particular program set (e.g., the measure could be used across programs or is included in a MAP “family of measures”); or • The value to patients/consumers outweighs any burden of implementation. 	<p>Yes: Review can continue</p> <p>No: Highest rating can be Refine and Resubmit.</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>

Assessment	Definition	Outcome
5. The measure can be feasibly reported.	<ul style="list-style-type: none"> The measure can be operationalized (e.g., the measure is fully specified; specifications use data found in structured data fields; and data are captured before, during, or after the course of care). 	<p>Yes: Review can continue</p> <p>No: Highest rating can be Refine and Resubmit.</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>
6. The measure is reliable and valid for the level of analysis, program, and/or setting(s) for which it is being considered.	<ul style="list-style-type: none"> The measure is NQF-endorsed; or The measure is fully developed and full specifications are provided; and Measure testing has demonstrated reliability and validity for the level of analysis, program, and/or setting(s) for which it is being considered. 	<p>Yes: Measure can be supported or conditionally supported.</p> <p>No: Highest rating can be Refine and Resubmit.</p> <p>MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>
7. If a measure is in current use, no unreasonable implementation issues that outweigh the benefits of the measure have been identified.	<ul style="list-style-type: none"> Feedback from end users has not identified any unreasonable implementation issues that outweigh the benefits of the measure; or Feedback from implementers or end users has not identified any negative unintended consequences (e.g., premature discharges, overuse or inappropriate use of care or treatment, limiting access to care); and Feedback is supported by empirical evidence. 	<p>If no implementation issues have been identified: Measure can be supported or conditionally supported.</p> <p>If implementation issues are identified: The highest rating can be Conditional Support. MAP can also choose to not support the measure, or request it be revised and resubmitted. MAP will provide a rationale for the decision to not support or make suggestions on how to improve the measure for a future support categorization.</p>

Improvements to MAP/CDP Integration

NQF made a number of improvements to integrate the MAP and CDP processes to provide MAP members and the public better information about the endorsement status of measures under consideration. Results of the CDP review of measures under consideration are included in the preliminary analysis and in the discussion guide provided to MAP members. NQF staff now reaches out to developers of measures given conditional support pending NQF endorsement to let them know of opportunities to submit their measure for review. MAP recommendations are provided to the relevant CDP standing committee when a measure

is reviewed for endorsement. Finally, MAP may refer specific questions about endorsed measures to be considered by the CDP committee.

Review of Current Program Measure Sets

The annual review of the measures under consideration provides measure-by-measure input. However, MAP's Measure Selection Criteria emphasizes the characteristics of an ideal program measure set. MAP expressed a need to better understand the program measure sets in their totality. Specifically, MAP members would like to understand how the measures under consideration would interact with the current measures in the

program set, the NQF endorsement status of current measures, and any end-user experience with current measures. As an improvement to the pre-rulemaking process in 2016-2017, MAP conducted a review of current program measure sets, offering guidance and input on ways to strengthen current measure sets by program, including recommendations for future removal of measures. This guidance, along with measure gaps across settings, is built into the final MAP reports, but not reflected in the spreadsheet of MAP final recommendations. Potential criteria for measure removal from a program measure set included:

- The measure is not evidence-based and not linked strongly to outcomes.
- The measure does not address a quality challenge (i.e., measure is topped out).
- The measure does not use measurement resources efficiently or contributes to misalignment.
- The measure cannot be feasibly reported.
- The measure is not NQF-endorsed or is being used in a manner inconsistent with NQF endorsement.
- The measure has lost NQF endorsement.
- Unreasonable implementation issues that outweigh the benefits of the measure have been identified.
- The measure may cause negative unintended consequences.
- The measure does not demonstrate progress toward achieving the goal of high-quality, efficient healthcare.

Feedback Loop Pilot

MAP members have expressed interest in knowing more about what happens to a measure after MAP has reviewed it. MAP members want to know if a measure has been submitted for NQF endorsement and the results of the endorsement review. Similarly, MAP wants to know whether a developer has updated a measure to address improvements suggested by MAP. Finally, MAP wants to know if a measure is performing as expected.

For the 2016-2017 pre-rulemaking cycle, NQF and CMS piloted a “feedback loop” process with the Post-Acute Care/Long Term Care Workgroup. During the Workgroup’s web meeting, NQF and CMS provided updates on the development and endorsement of selected measures. The goal of the feedback loop was to provide updates based on stakeholder concerns. This review was not intended to allow for a change in MAP’s recommendation about a measure.

The feedback loop was well-received by MAP members and the public. Based on the success of this pilot, NQF will work with CMS and other measure stewards and developers to provide updates on relevant measures across MAP. As noted in the decision categories above, the feedback loop will provide an opportunity for MAP members to better understand how their suggested refinements and conditions of support have been met.

BACKGROUND ON RECOMMENDATIONS

As noted above, MAP reaches a decision about every measure under consideration. MAP’s recommendations on individual measures for particular programs are found in the [spreadsheet of MAP final recommendations](#). Each decision is accompanied by one or more statements of rationale that explain why the decision was reached. Table 5 outlines the decision categories along with sample rationales for each category.

Voting Process

MAP members who identified selected measures for discussion described their perspective on

the use of the measure in the federal program, and how their perspective differed from the preliminary recommendation provided in the discussion guide. Other MAP members participated in the discussion to share their opinions. After discussion, MAP members voted on the measure under consideration. For the 2016-2017 pre-rulemaking cycle, MAP members had four voting options: (1) support for rulemaking, (2) conditional support for rulemaking, (3) refine and resubmit prior to rulemaking, and (4) do not support for rulemaking. MAP defined consensus as a vote of 60 percent or greater on each measure under consideration (Table 6).

TABLE 5. MAP DECISION CATEGORIES AND EXAMPLE RATIONALES

MAP Decision Category	Example Rationales
Support for Rulemaking	<ul style="list-style-type: none"> • Meets a critical program objective • Addresses a previously identified measure gap • Core measure not currently included in the program measure set • Promotes alignment across programs and settings
Conditional Support for Rulemaking	<ul style="list-style-type: none"> • Not ready for implementation; should be submitted for and receive NQF endorsement
Refine and Resubmit Prior to Rulemaking	<ul style="list-style-type: none"> • Not ready for implementation; measure needs further experience or testing before being used in the program
Do not Support for Rulemaking	<ul style="list-style-type: none"> • Overlaps with a previously finalized measure • A different NQF-endorsed measure better addresses the needs of the program • Does not meet a critical program objective

TABLE 6. VOTING PATHWAY

Do Not Support	Refine and Resubmit	Conditional Support	Support
>60% threshold of do not support OR <60% threshold for the combined total of refine and resubmit, conditional support, and support	≥60% threshold of refine and resubmit OR ≥60% threshold of refine and resubmit, conditional support, and support	≥60% threshold of conditional support OR ≥60% threshold of conditional support and support	≥60% threshold of support

NATIONAL QUALITY FORUM
1030 15TH STREET, NW, SUITE 800
WASHINGTON, DC 20005

www.qualityforum.org