MEASURE APPLICATIONS PARTNERSHIP

Process and Approach for MAP Pre-Rulemaking Deliberations, 2015-2016

FINAL REPORT

FEBRUARY 2016

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

- There were process improvements to MAP this year, including the addition of a one-day in-person meeting for the MAP Coordinating Committee, refinements to the preliminary analysis of measures, and updates to the consensus building and voting process.
- This year, MAP examined 131 unique measures for potential use in 16 different federal health programs.

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 Medicaid & Medicare 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 decisionmaking and provides specific recommendations about the best use of available measures and filling measure gaps.

PROCESS AND APPROACH

Overall Approach

During the 2015-2016 pre-rulemaking period, MAP used a three-step process to analyze and select measures.

1. **Develop Program Measure Set Framework**. Using CMS critical program objectives and NQF Measure Selection Criteria, NQF staff developed a framework for each program measure set in order to organize each program's finalized measure set. These frameworks were used to better understand the current measures in the program and how well any new measures might fit into the program by allowing workgroup members to quickly identify gaps and other areas of need.

2. **Evaluate measures under consideration for what they would add to the program measure sets**. 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 based on the measures under consideration based on the MAP-approved preliminary analysis guidance, included in Appendix A. The MAP workgroups discussed their recommendations for each measure under consideration during December in-person meetings.
3. **Identify and prioritize gaps for programs and settings.** MAP continued to identify gaps in measures within each program and provided measure ideas to spur development. MAP also considered the gaps across settings, prioritizing by importance and feasibility of addressing the gap when possible.

Review of 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.

Review of Specific Measures Under Consideration

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 131 unique measures for potential inclusion in 16 federal health programs. Since some measures were considered for multiple programs, MAP made 141 recommendations on measures within in a particular program.[a] 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 formalized process in which stakeholders can provide feedback on individual measures immediately after HHS provides the year's list of measures under consideration. These public comments were taken into account when MAP workgroups reviewed the measures under consideration in December. There was another opportunity for public comment in which 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 decisions on measures to be included in the quality reporting programs. Furthermore, both NQF members and any interested parties of 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. 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 list received on November 25, 2015, contained 131 unique measures for 16 different federal health programs. As some measures were considered for multiple programs, the list described 141 different situations where a particular measure could be selected for a particular program.

IMPROVEMENTS THIS YEAR

In an effort to continually improve and refine the MAP process, NQF used feedback from external stakeholders, MAP members, and NQF members to build upon changes made last year. 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.

MAP Coordinating Committee Fall In-Person Meeting

Based on recent feedback, NQF recognized a need to have the Coordinating Committee take the lead in framing the strategic direction for the overall MAP process and provide stronger guidance to the MAP workgroups. To achieve this goal, NQF hosted a one-day in-person meeting with the Coordinating Committee in the fall. During the meeting, the Coordinating Committee provided upstream guidance to the MAP workgroups on how to consider the measures under consideration. The Coordinating Committee provided input on reviewing the MUCs for "potential and actual impact" within programs as well as guidance on defining gaps in program measure sets. Ultimately, this meeting led to refinements to the MAP preliminary analysis process. Based on the discussion at the meeting, NQF recognized that MAP may need to create a core concept measurement framework that would represent the aspirational measurement goal of identifying gap areas among measures across programs and settings under the pre-rulemaking task.

All-MAP Meeting

Prior to the workgroups' deliberations, NQF hosted an all-MAP web meeting in the fall. This meeting provided an overview of the pre-rulemaking approach and the mechanism for reviewing the measures under consideration during the pre-rulemaking cycle. The goal was to ensure that all workgroup members had a clear understanding of their charge prior to attending the December in-person meetings. The web meeting included an orientation from CMS on how the MUC list is created, a review of the pre-rulemaking approach, including the preliminary analysis algorithm, the voting process, and the discussion guide tool used by the workgroups during their in-person meetings.

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):

- Does the measure under consideration meet a critical program objective as defined by MAP?
- Is the measure under consideration fully developed?
- Is the measure under consideration tested for the appropriate setting and/or level of analysis for the program? If no, could the measure be adjusted to use in the program's setting or level of analysis?
- Is the measure under consideration currently in use? If yes, does a review of its performance history raise any red flags?

- Does the measure under consideration contribute to the efficient use of measurement resources for data collection and reporting and support alignment across programs?
- Is the measure under consideration NQF-endorsed for the program's setting and level of analysis?

For measures in an early stage of development, MAP may not have the information to answer all of the questions listed above. In addition, early-stage measures may change as they undergo testing and further development. Therefore, MAP evaluated these measures using an abbreviated algorithm, which sought to encourage the development of innovative new measures while maintaining rigor. This is intended to provide CMS and measure developers with upstream information on the further development and potential applications for these measures. The preliminary analysis algorithm asks:

- Does the measure under consideration meet a critical program objective as defined by MAP?
- Is the measure under consideration fully specified?
- Does the measure under consideration contribute to the efficient use of measurement resources and support alignment across programs?

Refinements to the Consent Calendar

The measures under consideration were presented to the workgroups in a consent calendar format that grouped together similar measures. NQF staff presented the consent calendar reflecting the result of the preliminary analysis using the MAP selection criteria and programmatic objectives. Members were able to identify specific measures from the consent calendar that required further discussion. This approach allowed the groups to spend more time on measures where there were differing stakeholder perspectives and less time on measures where consensus had been reached. Once all measures that the workgroup wanted to discuss were removed from the consent calendar, the co-chairs asked if there was any objection to accepting the preliminary recommendation provided by NQF staff for the MUCs remaining on the consent calendar. When no objections were raised, the consent calendar and the associated preliminary recommendations were accepted. However, when objections were raised, the workgroup discussed these objections and determined whether it warranted removing measures from the consent calendar for individual vote or if the workgroup simply needed clarification about the preliminary recommendations.

Voting Process

Participants who identified the need for discussion on select measures described their perspective on the use of the measure and how it differed from the preliminary recommendation provided in the discussion guide. Workgroup members assigned as lead discussants for the relevant group of measures were asked to respond to the individuals who requested discussion. Lead discussants stated their own point of view, whether it agreed with the preliminary recommendation or the divergent opinion. Other workgroup members participated in the discussion to share their opinions. ASfter discussion of each MUC, the workgroup used two voting pathways: for fully developed measures and measures under development (see tables 1 and 2 below).

To establish consensus, a vote of 60 percent or greater needed to be reached on each measure under consideration.

TABLE 1. FULLY DEVELOPED MEASURES VOTING PATHWAY

Do not support	Conditional support	Support
>60% consensus of do not support OR <60% consensus for the combined total of conditional support and support	≥60% consensus of conditional support OR ≥60% consensus of both conditional support and support	≥60% consensus of support

TABLE 2. MEASURES UNDER DEVELOPMENT VOTING PATHWAY

Encourage continued development	Do not encourage further consideration	Insufficient information
≥60% consensus of encourage continued development	≥60% consensus of do not consider further consideration	≥60% insufficient information

BACKGROUND ON RECOMMENDATIONS

MAP's recommendations on individual measures for particular programs are provided in an accompanying spreadsheet. Each decision is accompanied by one or more statements of rationale that explain why the decision was reached. Tables 3 and 4 outline the recommendation categories along with sample rationales for each category.

TABLE 3. MAP DECISION CATEGORIES AND EXAMPLE RATIONALES FOR DEVELOPED MEASURES

MAP Decision Category	Example Rationales
Support	• 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	• Not ready for implementation; should be submitted for and receive NQF endorsement • Not ready for implementation; measure needs further experience or testing before being used in the program.
Do not support	• 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 4. MAP DECISION CATEGORIES AND EXAMPLE RATIONALES FOR EARLY-STAGE MEASURE DEVELOPMENT

MAP Decision Category	Example Rationales
Encourage continued development	• Addresses a critical program objective, and the measure is in an earlier stage of development • Promotes alignment, and the measure is in an earlier stage of development
Do not encourage further consideration	• Overlaps with a finalized measure for the program, and the measure is in an earlier stage of development. • Does not address a critical objective for the program, and the measure is in an earlier stage of development.
Insufficient Information	• Measure numerator/denominator not provided

APPENDIX A:
Preliminary Analysis Guidance for Measures Under Consideration (MUC)

Part 1 – Developing a Program Measure Set Framework

- Using the critical program objectives, NQF staff develops an organizing structure or framework for each program measure set. The framework builds off of the National Quality Strategy aims and priorities.
- Using the framework, NQF staff organizes the measures currently finalized in the program.

Part 2 – Completing Preliminary Analysis of MUCs

NQF staff then evaluates each MUC within the context of the program measure set to see what it might add to the program. Staff answers a series of questions for all MUCs to help inform MAP Workgroup members during their deliberations. The preliminary analysis uses the MAP Measure Selection Criteria and the CMS 2015 Measures under Consideration List Program Specific Measure Priorities and Needs.

Step 1: Does the MUC meet the program goals and objectives?

Refer to *MAP MSC #3* "Program measure set is responsive to specific program goals and requirements" **and** *CMS MUC Measure Selection Requirement (MSR) 2a* "Measure is responsive to specific program goals and statutory requirements."

- Using the CMS 2015 Program Specific Measure Priorities and Needs: determine how/whether the MUC addresses the program goals and objectives.
- How does the MUC address specific program objectives and measurement requirements that are not already addressed by existing measures?
- If the measure does not address a critical program objective, MUC receives a **Do not support** for its preliminary analysis.

Step 2: Is the MUC a high-value measure?

High-value measures are defined as "measures that will drive the healthcare system to higher performance." Refer to *CMS MSR 2b* "Measure addresses an important condition/topic with a performance gap and has a strong scientific evidence base to demonstrate that the measure when implemented can lead to desired outcomes and/or more affordable care."

MAP has identified the following measure types as high-value:

- Outcome measures (e.g., mortality, adverse events, functional status, patient safety, complications, or intermediate outcomes, e.g., BP value, lab test value—not just that the test is performed)
- Patient-reported outcomes where the patient provides the data about their results of treatment, level of function, and health status (NOT the clinician administering a tool/questionnaire for the patient to fill out—the measure must use the results of the information in the tool or questionnaire)
- Measures addressing patient experience, care coordination, population health, quality of life, or impact on equity. *MAP MSC #5 and #6*
- Appropriateness, overuse, efficiency, and cost-of-care measures
- Composite measures
- Process measures close to outcomes with a strong evidence link.

Step 3: Does the MUC fill a gap in the program measure set?

- Is the measure in a MAP Family of Measures? Families of measures provide a tool that stakeholders can use to identify the most relevant available measures for particular measurement needs; to promote uniformity by highlighting important measurement categories; and to apply to other measurement initiatives. With its 2014 report, MAP has now produced 10 families that assess all 6 priorities within the National Quality Strategy.
- Does it address a high priority domain identified by CMS that does not have adequate measures in the program set?
- If the measure does not fill a gap, MUC receives a **Do not support** for its preliminary analysis.

Step 4: Is the MUC fully specified?

NQF staff uses the definition of fully specified in *CMS MSR 2e* "Measure reporting is feasible and measures have been fully developed and tested. In essence, measures must be tested for reliability and validity."

- If the measure development status on the MUC list is "early development" or "field testing," the MUC is not fully developed and will be evaluated using the "Measure Under Development Pathway" (see below)
- If the MUC is fully specified and tested, it will be evaluated as a fully developed measure using the steps below.

Step 5: Is the MUC tested for the appropriate setting and/or level of analysis for the program?

If the measure is specified and tested for a different setting or level of analysis that is not appropriate for this program (e.g., a MUC for clinician programs that is specified/tested/endorsed at the health plan level only):

- Hospital measures receive a **Do not support** for their preliminary analysis.
- PAC/LTC measures: Could a hospital measure be used in the PAC/LTC setting or be "tweaked" to use in the PAC/LTC setting? If yes, continue on to step, 6 but note that any support must be conditional on the measure being tested with PAC/LTCs before being used in a public reporting or payment program. If no, PAC/LTC measures receive a **Do not support** for their preliminary analysis.
- Clinician measures: Could the measure be used at the clinician level or be "tweaked" to use at the clinician level? If yes, continue on to step 6, but note that any support must be conditional on the measure being tested at the clinician level before being used in a public reporting or payment program. If no, Clinician measures receive a **Do not support** for their preliminary analysis.

Step 6: Is the MUC currently in use? If not in use, go to step 7.

- Determine if the MUC is currently in use in another federal quality initiative program or in a private sector program such as ones sponsored by health plans and purchasers. The MUC list generally indicates use in other programs.
- If in use, staff searches out any information on measure performance using the following sources:
  - Public reports
  - CMS or HHS reports (Impact report; National Quality Report, etc.)
  - Public reporting websites
  - Information from OPUS if measure has been submitted for endorsement
  - Dry run data for some programs
  - Information from the developers
  - PubMed

- If no performance data are identified, that will be indicated in the preliminary analysis.
- Staff looks for any "red flags":
  - What is current performance? Is the measure performance close to 100 percent, i.e., is it topped out?
  - Is there a history of implementation challenges (e.g., data source issues)?
  - Does the measure lead to misalignment (if information on specification is available)?
  - Are there any known unintended consequences?
  - Does the measure have a low selection rate among providers (for PQRS measures)?
    - » PQRS utilization report.

Step 7: Does the MUC contribute to alignment and efficient use of measurement resources (burden and cost of measurement)? *MAP MSC #2-7*

- Is the measure used in other programs?
- Is this the best measure available (e.g., outcome measures are preferred over process measures)?
- Not duplicative of an existing measure BUT also consider whether the MUC is a better measure
  - If a MUC is thought to be a better measure, it receives **Conditional support**—conditional on replacing the existing measure
- Captures the broadest population
- If the topic area already has outcome measures, is this process measure needed?
- Composite measures
- Consider the cost-benefit balance. Weigh the burden of implementation against the value of the measures for patients (e.g., implementing PROs may be burdensome but is extremely high-value).
- If the measure does not contribute to the efficient use of resources or support alignment across programs, the MUC receives a **Do not support** for its preliminary analysis. If yes, go to step 8.

Step 8: NQF endorsement status

MAP MSC # 1 "NQF-endorsed measures are required for program measure sets, unless no relevant endorsed measures are available to achieve a critical program objective."

- NQF-endorsed, or likely to receive NQF-endorsement in the near future at the level of analysis and for the setting in the program: MUC to receive a **Support** for its preliminary analysis
- Never submitted for NQF endorsement; OR failed initial endorsement submission but has since been modified to reflect NQF CDP Standing Committee feedback; OR a measure not specified at the clinician level that could be used at the clinician-level: measure receives **Conditional support** for its preliminary analysis. State condition that must be met and conditionally support MUC.

  Previous examples of conditions include, but are not limited to:
  - Not ready for implementation; should be updated to reflect current guidelines.
  - Not ready for implementation; data sources do not align with program's data sources.
  - Not ready for implementation; should be submitted for and receive NQF endorsement.
  - The measure must be tested at the clinician level before being used in a public reporting or payment program.
  - Better measure to replace existing measure.

- Submitted for NQF endorsement, but not recommended by NQF CDP Standing Committee: MUC to receive a **Do not support** for its preliminary analysis.

Summary of MAP Recommendations for Fully Developed Measures

- **Support** measures that meet all the criteria above and are NQF-endorsed.
- **Conditionally support** measures that may have an issue that could be addressed as a condition.
- **Do not support** measures identified during the preliminary analysis above.

Measures Under Development Pathway

This pathway is used for measures that have met the critical program objectives, and the high-value and gap criteria but are not yet fully developed.

Does the MUC contribute to the efficient use of resources and support alignment across programs?

Staff refers to *MAP MSC 2-7* to answer this question.

A measure would support alignment if:

- Not duplicative of a measure currently in the program or would be a better measure
- Aligns with similar measures in the same program or other programs

A measure demonstrates efficient use of measurement resources (burden and cost of measures) if:

- Captures the broader population
- High-value measure – see above
- If the topic area has outcome measures, is this process measure needed?
- Composite measures

If the measure under development meets criteria for alignment and efficient use of resources, the measure receives **Encourage continued development** for its preliminary analysis.

If the measure under development does not contribute to the efficient use of resources or support alignment across programs, the measure receives **Do not encourage further consideration** for its preliminary analysis.

Summary of MAP Recommendations for Measures Under Development

- **Encourage continued development** of measures that meet above the criteria.
- **Do not encourage further consideration** of measures that do not contribute to efficient use of resources or alignment across program.
- **Insufficient information** for measures with limited to no information specified.

ISBN 978-1-68248-006-9

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

www.qualityforum.org