Impact Assessment of CMS Quality and Efficiency Measures
Contract #: 75CMC18D0026; Task Order #: 75FCMC190001

CMS Quality Measure Development Plan/ Quality Measure Index Technical Expert Panel Meeting Summary

Meeting Date: December 17, 2021

Submitted: January 4, 2022

To: Nidhi Singh Shah, Project Lead
HHS/CMS/OA/CCSQ/QMVIG/DPMS
7501 Wisconsin Ave., Mail Stop 8-008
Bethesda, MD 20814
Nidhi.Singh-Shah@cms.hhs.gov

Erika Armstrong, Contracting Officer's Representative (COR)
HHS/CMS/OA/CCSQ/QMVIG/DPMS
Centers for Medicare & Medicaid Services (CMS)
7500 Security Boulevard, Mail Stop S3-02-01
Baltimore, MD 21244-1850
Erika.Armstrong@cms.hhs.gov

HSAG HEALTH SERVICES ADVISORY GROUP

Table of Contents

I. Introduction 1
II. Meeting Proceedings 1
Welcome and Opening Remarks 1
TEP Roll Call and Disclosures of Conflict of Interest 1
Future Plans for the Measure Development Plan (MDP) Project 2
2022 MDP Annual Report Objectives 2
2022 Environmental Scan & Gap Analysis Report 2
Updates for QMI Project: Development Status and Lessons Learned 3
Purpose of the QMI 3
QMI Structure 3
QMI Scope 3
QMI Completed Milestones 3
Lessons Learned From QMI Testing of Existing Measures 4
Recommendations From Stakeholders 4
Standardized Pre-Rulemaking Measure Data Submission 4
Updates for QMI Project: Review of Updated Methodology 4
QMI Scoring Approach 4
Refined Variable Scoring 5
Importance Domain 5
Evidence 5
High Priority 5
Measure Performance 5
Scientific Acceptability Domain 6
Reliability 6
Validity 6
Reliability and *Validity* Scoring Approach 6
Risk Adjustment 7
Feasibility & Usability Domain 7
Provider Burden 7
Discussion 7
Updates for QMI Project: Future Enhancements to QMI 9
Future QMI Development 9
Discussion 9
Next Steps 11

Technical Expert Panel Meeting Summary

I. Introduction

The Centers for Medicare & Medicaid Services (CMS) has contracted with Health Services Advisory Group, Inc. (HSAG) to develop the CMS Quality Measure Development Plan: Supporting the Transition to the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs) under Contract #75FCMC18D0026; Task Order #75FCMC19F0001. As part of this contract, HSAG ("the team") is also tasked to develop the CMS Quality Measure Index (QMI).

As part of this contract, HSAG convenes a Measure Development Plan/Quality Measure Index technical expert panel (MDP QMI TEP) of patients and family caregivers, clinicians and representatives of professional societies, consumer advocates, quality measurement experts, and health information technology specialists to provide multi-stakeholder input on project tasks and reports. A workgroup within the MDP QMI TEP provided feedback on development of the QMI.

On December 17, 2021, the team convened a webinar meeting of the TEP to discuss upcoming objectives for the MDP project and updates to the QMI based on input from the TEP workgroup.

II. Meeting Proceedings

Welcome and Opening Remarks

Presenter: Michelle Pleasant, PhD, MA; HSAG

Dr. Pleasant, HSAG lead for the MDP project, welcomed TEP members. After reviewing standard ground rules for the meeting, she presented the meeting agenda, which included:

- Upcoming objectives for the MDP project
- Summary of TEP workgroup input and corresponding updates to QMI scoring variables
- Future plans for the QMI project
- 2022 MDP QMI Call for TEP Nomination Period

TEP Roll Call and Disclosures of Conflict of Interest

Presenter: Amy Mullins, MD, CPE, FAAFP, TEP Chair

Dr. Mullins, TEP chair, conducted the roll call. She noted that TEP members with asterisks next to their names on the roster were members of the TEP workgroup. Thirteen of 17 TEP members were present.

TEP	CMS (optional)	HSAG
☒ Anders, Scott*	☐ Armstrong, Erika	☒ Campbell, Kyle
☐ Aran, Peter	☒ Dollar-Maples, Helen	☒ Hanley, Kendra
☒ Bossley, Heidi*	☒ Singh Shah, Nidhi	☒ Hall, Marie
☒ Fields, Robert		☒ Hemmingway, Susan
☒ Huang, Mark		☒ Keenan, Megan
☐ Kaufman, Joel		☒ Lockwood, Carolyn
☒ Mosnaim, Giselle		☒ Mackeprang, Julia
☒ Mullen, Cody		☒ Nguyen, Kim

TEP	CMS (optional)	HSAG
☒ Mullins, Amy ☒ Nguyen Howell, Amy ☐ Nielsen, Matthew ☒ Rising, Kristin ☒ Rogut, Lynn ☒ Scholle, Sarah* ☒ Suter, Lisa Gale* ☒ Tierney, Samantha* ☒ Wisham, Lindsey*		☒ Pleasant, Michelle ☒ Selvarajah, Shalini ☒ Yang, Sherry ☒ Ziemba, Rob

*TEP workgroup member

The following TEP members provided status updates and/or disclosed conflicts of interest:

- Giselle Mosnaim stated her correct organizational affiliation is the American Academy of Allergy, Asthma, and Immunology, which nominated her for the TEP. She disclosed that she has research grants from Teva Pharmaceuticals, Sanofi Pharmaceuticals, Regeneron Pharmaceuticals, and GlaxoSmithKline.
- Cody Mullen noted he is now with Purdue University as a clinical associate professor and has been appointed to the National Quality Forum (NQF) Rural Health Advisory Group.
- Kristin Rising informed that she is the director of the Center for Connected Care at Thomas Jefferson University.
- Lynn Rogut stated she is now retired.
- Sarah Scholle noted she is with the National Committee for Quality Assurance (NCQA), adding this is her only conflict.
- Lindsey Wisham commented that her employer is Telligen, which has CMS contracts; however, she serves on the TEP as a patient representative.
- Lisa Gale Suter disclosed her salary support is through a federal contract from CMS to develop measures; however, to her knowledge none are under consideration by the TEP.

Future Plans for the Measure Development Plan (MDP) Project

Presenter: Michelle Pleasant, PhD, MA; HSAG

Dr. Pleasant presented an update of the MDP project, stating the team recently completed the first draft of the 2022 MDP Annual Report. After the CMS clearance process, which will begin in spring 2022, the final report will be posted on the CMS MDP website May 1, 2022. Required by statute, the report provides updates on CMS progress in achieving objectives of the MDP. The report tracks the status of measure gaps, developed measures, and measures in development, and provides an inventory of measures in the Quality Payment Program.

2022 MDP Annual Report Objectives

The 2022 MDP Annual Report will ask stakeholders and CMS colleagues how their measure development addresses health equity and how the ongoing COVID-19 public health emergency has continued to affect measure development activities.

2022 Environmental Scan and Gap Analysis Report

Dr. Pleasant noted work on the 2022 environmental scan and gap analysis will begin this spring. The report will focus on CMS-prioritized clinical areas for future Merit-based Incentive Payment System (MIPS) Value Pathway (MVP) development. Currently, seven MVPs are included in the 2022 Physician Fee Schedule Final Rule. These

include rheumatology, stroke care, heart disease, chronic disease management, emergency medicine, lower extremity joint repair, and anesthesia.

The environmental scan will focus on clinical areas prioritized by CMS for future MVP development, which are behavioral health, kidney disease, hypertension, maternal and women's healthcare, diabetes, and human immunodeficiency virus (HIV). The scan will include an inventory of existing MIPS measures to determine what is currently available for MVP development, and the gap analysis to guide future measure development.

Updates for QMI Project: Development Status and Lessons Learned

Presenter: Carolyn Lockwood, MSN, RN; HSAG

Carolyn Lockwood, QMI project lead, indicated the purpose of her presentation would be to discuss the QMI and refinements made based on input from TEP members.

Purpose of the QMI

Ms. Lockwood described the QMI as an index that can be used to assess the relative value of quality measures based on key measure characteristics. The QMI can assist CMS in addressing a recommendation by the U.S. Government Accountability Office (GAO) for systematic measure assessment. Ms. Lockwood noted that not all measures undergo other expert review processes prior to CMS submission and emphasized the QMI is not intended to replace existing expert review processes but rather to complement them. Ms. Lockwood added that the QMI can assist CMS in prioritizing measures and can be adapted across programs, the Measure Lifecycle, and as measure priorities change.

QMI Structure

Ms. Lockwood indicated the QMI is structured by operationalizing key measure characteristics into two types of variables:

- *Classification variables* can be used to group or stratify the measures being assessed. These variables are not factored into the QMI score.
- *Scoring variables* capture the information integral to the assessment of measures and are organized into three domains:
    - Importance
    - Scientific Acceptability
    - Feasibility & Usability

  There are currently eight scoring variables in the QMI.

QMI Scope

To date the QMI has been tested using measures with the following characteristics:

- Outcome, intermediate outcome, PRO-PM, and process measure types
- Clinician-, group-, and facility-level attribution

Ms. Lockwood added that future considerations for testing the QMI could encompass the following:

- Structure, cost/resource use, and patient experience measures
- Health plan- and population-level attribution

QMI Completed Milestones

Ms. Lockwood reviewed the QMI development milestones the team has completed thus far:

- Developed and conducted three comprehensive environmental scans to identify and confirm variables.
- Obtained multi-stakeholder feedback, including from CMS, the TEP, and the TEP workgroups.
    - Ms. Lockwood thanked members of the TEP who participated in the workgroup and indicated that later in the presentation the HSAG team would review revisions made to the QMI based on workgroup recommendations.
- Assessed 207 unique measures in the 2020 Quality Payment Program.
- Conducted testing to adapt the QMI for broader use in facility-level measures across CMS programs.
- Reviewed measure data collected through the 2021 pre-rulemaking cycle and standardized data collection for the 2022 pre-rulemaking cycle.

Lessons Learned from QMI Testing of Existing Measures

Ms. Lockwood reviewed lessons learned from QMI testing of existing measures; specifically:

- Measure information provided by stakeholders is often imprecise and heterogenous.
- These factors limit CMS in conducting a fair comparison of the relative value of quality measures in achieving CMS objectives and selecting measures for use in their programs.
- Measure developers often provide incomplete information related to the measure testing.

Ms. Lockwood indicated NQF has identified similar issues in its review process and is working to clarify accepted analyses and corresponding thresholds. The QMI will align with potential NQF recommendations to the extent feasible.

Recommendations from Stakeholders

Ms. Lockwood acknowledged the TEP workgroup had previously indicated a preference for simpler QMI scoring algorithms and agreed with refinements to key variables and variable components, including:

- *Evidence*: score based on best evidence provided
- *Reliability* and *Validity*: score based on best result when multiple accepted analyses are provided
- *Reliability* and *Validity*: Survey-level testing – separate from performance measure-level testing components of the variables.

The workgroup also recommended removing the social risk factor requirement from the *Risk Adjustment* variable to provide flexibility.

Standardized Pre-Rulemaking Measure Data Submission

Ms. Lockwood reviewed HSAG activities to provide input to refine and standardize data collection during the annual pre-rulemaking cycles. HSAG is working with CMS to provide input on the CMS Measures under Consideration Entry/Review Information Tool (MERIT) for collecting information on measures during pre-rulemaking. The goal of this work is to provide more discrete fields to standardize data collection to allow a fairer comparison of measures. These activities are laying the groundwork for future automation of the QMI.

Updates for QMI Project: Review of Updated Methodology

Presenter: Rob Ziemba, PhD; HSAG

Next, Dr. Rob. Ziemba presented a review of updated methodology for the QMI.

QMI Scoring Approach

Dr. Ziemba discussed the three domains and the variables included in each domain:

- Importance

  - *Evidence*
  - *High Priority*
  - *Measure Performance*
- Scientific Acceptability
  - *Reliability*
  - *Validity*
  - *Risk Adjustment*
- Feasibility and Usability
  - *Provider Burden*
  - *Feasibility*

Dr. Ziemba noted the variables within each domain are averaged to obtain the domain score and that the domains are currently equally weighted to obtain the overall QMI score. The scores for domains and the overall score are scaled from 0 to 100.

Refined Variable Scoring

Dr. Ziemba provided more detail on how each variable is scored. All variables are scored based on four color-coded criteria: Green – Preferred (1.00), Yellow – Acceptable (0.75), Red – Not preferred (0.25), and Grey - Missing (0.00). He explained that the QMI uses CMS and other industry standards and thresholds to score variables where possible. He added that use of consistent categories across variables allows for easier comparison and interpretability of scores.

Importance Domain

Dr. Ziemba reviewed the definitions for each of the three variables in the Importance domain: *Evidence*, *High Priority*, and *Measure Performance*. This domain aims to assess how well a measure is supported by evidence, whether it addresses a current CMS priority, and whether there is a measurable gap in quality.

Evidence

Dr. Ziemba reviewed the hierarchy of evidence and noted the QMI scores for the variable are based on the best piece of evidence provided. He indicated the QMI scoring has been simplified based on workgroup feedback. The QMI places the quality of the evidence into broad categories and assumes the evidence provided by the developer is appropriate to support the measure. He noted that the QMI is not meant to replace review of the provided evidence's quality by CMS, NQF, or other parties.

High Priority

Dr. Ziemba noted there are currently three CMS priorities for this variable and that a measure would score Green if it met two or more priorities, Yellow if it met one priority, and Red if it did not meet any CMS priority.

Measure Performance

Dr. Ziemba noted there is no universal definition of benchmarks for quality measures across programs. From the distribution of measure scores, the QMI assigns the maximum as the benchmark for measures where a higher score is better and the minimum for measures where a lower score is better. When not available for proportion measures, the QMI imputes 100 or 0 as the benchmark depending on whether a higher score is better or a lower score is better. For non-proportion measures, this variable is scored as missing if information on the distribution for the benchmark is unavailable. This change was made based on workgroup feedback to use distributions where possible. Dr. Ziemba added that the *Variation* variable is no longer included in the QMI score because it also

reflected the distribution of scores and would be somewhat redundant with the *Measure Performance* variable as currently specified.

The *Measure Performance* variable is scored as the relative improvement needed to achieve the benchmark. Measures with room for improvement score Green while measures with very little room for improvement score Red. There is no Yellow category for this variable.

Scientific Acceptability Domain

Dr. Ziemba shared the three variables included in the Scientific Acceptability domain: *Reliability*, *Validity*, and *Risk Adjustment*. This domain aims to assess how likely the measure is to give similar results upon repeated measurements of the same unit, whether the measure captures the quality signal that it is meant to represent, and whether risk adjustment has been developed when appropriate. He noted there is a wide array of analyses that can be done to demonstrate these characteristics of measures; however, the field is converging on some accepted assessments and thresholds to standardize evaluation of the quality of submitted measures. The QMI has aligned with these where possible and will continue to evolve based on stakeholder input.

He added that the QMI is a rapid first order approximation of the scientific acceptability of measures. The QMI tracks to and aligns with NQF's thresholds where feasible. This approach allows for a low cost, rapid assessment of measures which could be confirmed by panels of experts who can make more nuanced judgments about the analyses submitted. He added that in alignment with NQF guidance, measure score level testing is preferred and results in higher QMI scores. However, the QMI accepts other levels of testing if measure score level testing is not provided.

Reliability

Dr. Ziemba reviewed the components of the *Reliability* variable: measure score, data element, and survey-level testing. Noting the thresholds and criteria to score each component, he pointed out that the variation and complexity of analyses provided increased for data element and survey-level testing. This posed a challenge in categorizing results, so the level of detail considered in scoring the components of reliability decreased for each of those levels. He provided the example of the heterogeneity in testing provided by developers to show the psychometric properties of surveys and noted that the QMI is only able to assess whether psychometric testing of a survey was done and not the quality of that testing.

Validity

Dr. Ziemba reviewed the components of the *Validity* variable: measure score, data element, face validity, and survey-level testing and the thresholds and criteria to score them. He noted this variable uses a similar approach as the *Reliability* variable. However, there was substantial heterogeneity in the analyses provided by developers to demonstrate measure score validity. Thus, the QMI only requires correlations or associations to be in the hypothesized direction and does not include thresholds that every measure should meet for those analyses.

Reliability and *Validity* Scoring Approach

Dr. Ziemba described how the *Reliability* and *Validity* variables use a hierarchical approach to scoring based on the types of analyses provided by the developer. For example, if measure score level reliability testing is provided, that determines the score for the *Reliability* variable. If it is not provided, the QMI would then consider whether data element level testing was conducted, and if not, whether survey-level testing was conducted to derive the variable score.

Risk Adjustment

Dr. Ziemba noted that the *Risk Adjustment* variable is only scored for outcome, patient-reported outcome-based performance measures (PRO-PMs), intermediate outcome, and resource use measures. He discussed that the team limited the measures scored on this variable based on feedback from the workgroup to not give credit to measures that do not need to be risk adjusted. He noted that by omitting this variable for process and structure measures, the other two variables in the Scientific Acceptability domain will weigh more than they will for measures scored on all three variables in the domain. However, this limitation is inherent in comparisons of measures with different properties.

Feasibility and Usability Domain

Dr. Ziemba described the two variables in this domain: *Provider Burden* and *Feasibility*. This domain aims to assess the degree to which a measure can be implemented in CMS programs with limited burden to providers. He noted these variables are based on current measures but may evolve as measure calculation and data collection methods change, such as introduction of the Fast Healthcare Interoperability Resources (FHIR) specification. He then described *Provider Burden* in more detail.

Provider Burden

Dr. Ziemba noted that the QMI addressed workgroup feedback by giving both claims measures and eCQMs the highest score for the *Provider Burden* variable (Green). Measures that require some effort from providers for calculation are scored Yellow and measures that must be calculated using manual abstraction, which requires the most effort from providers, are scored Red.

Discussion

- A TEP member noted that MIPS has other high priority categories such as "appropriate use" and "patient safety" that are not included in the current definition and asked whether those could fit into the QMI.
    - HSAG Response: The QMI has flexibility to adapt as CMS priorities shift. Further, the QMI is just one of many tools CMS can use in its decision-making process and CMS can always consider additional priorities beyond those in the QMI. "Appropriate use" and "safety" are part of the QMI classification variables; CMS could sort measures by those criteria even if they are not included in scoring.
- A TEP member asked how survey-based measures and PRO-PMs are scored in the Feasibility & Usability domain
    - HSAG Response: The domain is agnostic to measure type and focuses on the type of data and what effort is required of providers to calculate the measures. If a PRO-PM is specified as an eCQM, it would receive the top score for *Provider Burden* but if specified in a way that requires more effort from the provider to calculate the measure, it would receive a lower score according to the method of measure calculation.
    - The TEP member followed up by asking whether the QMI has considered patient experience measures or measures that use data outside of the provider's records.
    - HSAG Response: Patient experience measures were outside the scope of current QMI testing; the QMI may be adapted for those measure types in the future.
- A TEP member noted the discussed changes address much of the workgroup feedback. She asked how the QMI addresses measures that combine data sources, such as hybrid measures that use both claims and PRO-PM data.
    - HSAG Response: The QMI has considered hybrid measures and measures that allow for different measure calculation methods (e.g., eCQM and manual abstraction). The Yellow category for the

Provider Burden variable is a catch-all for everything that is not strictly claims or eCQM or strictly manual abstraction. Measures that are hybrid claims and eCQM would score in the Green category because both calculation methods are in that category.
    - A TEP member agreed the QMI should consider different scenarios for measures that allow for different calculation methods and for scenarios where more burdensome data collection is required to calculate PRO-PMs.
    - HSAG Response: The goal is eventually to obtain more granular data from developers on how the measures are expected to be calculated, with further refinement of the *Provider Burden* variable based on the additional information. The current approach is limited to assumptions based on data source and submission method because of the limited detail available in the current measure documentation related to provider burden.
- A TEP member noted signal-to-noise and test-retest produce different results for the same measures and indicated the QMI should continue to align with the evolving NQF Scientific Methods Panel (SMP) thresholds.
    - HSAG Response: It is correct that different analyses can produce different results, which can also differ depending on how developers compare rankings. The QMI team will continue to monitor the recommendations of the SMP and align the QMI where possible.
- A TEP member asked how often data are missing for variables and whether the lowest score when data are available for scoring should also be 0.
    - HSAG Response: The current definitions are based on measures currently in use. Data were often missing or not clearly specified in the publicly available forms for certain variables. The QMI evolved to score Red and Grey differently so that CMS can differentiate between poor performance and missing data. By adding more discrete fields to capture some of the information, the goal is to significantly reduce the amount of missing data for measures scored and thus be able to eliminate the missing fields and re-scale the variable scores.
- A TEP member asked whether *Reliability* and *Validity* of survey-based measures should score Yellow if only survey-level testing is provided. She noted that with this scoring, any survey-level testing is equivalent to an acceptable score for measure score-level testing and having good data element-level testing results. She suggested scoring survey-level testing lower since that type of testing does not assess the performance measure.
    - Two TEP members agreed that both performance measure and survey-level testing should be required for testing *Reliability* and *Validity*.
        - HSAG Response: The team agrees with both commenters. Scoring of the variables is hierarchical and if measure score-level testing is provided, that is used to score the QMI variable. The team is tracking how the NQF SMP is evolving on how to incorporate survey-level testing into the evaluation of measures and will align future versions of the QMI to the extent possible. The current version of QMI scoring for scientific acceptability was developed for existing measures in use in CMS programs and which were submitted to NQF in the past few years. In the review of existing measures, the team identified several developers that appeared to be unaware that both survey-level and performance measure-level testing were required. This may have occurred because the survey-level testing information is entered into the same fields in documentation as the performance measure testing results. The QMI team has suggested to NQF and CMS to clarify those documentation requirements and potentially create a separate section for survey-level testing. The goal would be to reconfigure the QMI scoring for these variables in the future to give lower scores to measures that only conducted survey-level

testing. The team did not feel it would be fair to penalize developers of current measures for this version of the QMI.
    - A TEP member asked a follow-up question about whether a Yellow for survey-level testing gets a 0.75 or a lower score because it is lower in the hierarchy.
        - HSAG Response: The only way to get the highest score for *Reliability* or *Validity* is to do measure score level testing. However, a Yellow for the other levels of testing would receive a score of 0.75, the same as would a Yellow for measure score level testing.
    - A TEP member added that the existing forms do have guidance to indicate that both survey-level and performance measure-level testing are required. She would not recommend including measures that have only been tested at the survey-level in accountability programs.
        - HSAG Response: The team agrees. The QMI has highlighted some issues with the current measure evaluation process, which potentially could be resolved by confirming the appropriate levels of testing have been conducted before implementing measures in CMS programs.
        - A TEP member added that it is possible a measure could be tested at the measure score level without testing at the survey level, which could also be an issue for the scientific acceptability of the measure.
- A TEP member asked why the scale does not include 0.5.
    - HSAG Response: The scale was intended to give higher scores to measures that were preferred or acceptable rather than scoring acceptable in the middle of the scale. The goal is that the color coding can be helpful for quickly seeing which variables are preferred, acceptable, and not preferred rather than relying on just the numbers.

Updates for QMI Project: Future Enhancements to QMI

Presenter: Carolyn Lockwood, MSN, RN; HSAG

Future QMI Development

Ms. Lockwood noted four areas for future QMI development:

1) Expand QMI measure assessment; to assess potential enhancements to the QMI, the team will review data submitted during the 2022 pre-rulemaking cycle.
2) Evaluate relevance of QMI variables to phases of the Measure Lifecycle and finalize criteria to apply to each phase of development.
3) Adapt the QMI for use among new measure types; the team will determine which variables should be updated to accommodate cost/resource use, patient experience, and structure measures, and also expand the types of levels of analysis.
4) Continue to align with updates to NQF testing requirements and CMS priorities as they evolve.

Discussion

Ms. Lockwood asked the TEP whether any additional concepts should be considered for future versions of the QMI.

- A TEP member asked whether an equity dimension had been considered. NQF has had disparity sensitive measures in the past. This differs from risk adjustment; it asks whether there are equity issues that would make it a priority.
    - HSAG Response: This is something the team is interested in and has reviewed, e.g., the NQF framework for disparities and sensitive measures. The challenge is defining the term – would this

be more of a classification variable or incorporated into scoring, and how would the data be standardized among developers?
    - The TEP member commented the disparity sensitive approach may belong in the Importance domain, where it is an opportunity. She offered to send HSAG a link to a recently published paper by her team related to Medicaid and thinking about an equity framework for measurement.
- A TEP member suggested considering stratification or equity as an adjunct criterion for *Risk Adjustment*. Another item to add are response rates for patient-reported measures, or lack of response, or response bias. She noted that as a developer of a PRO-PM, her team is struggling with balancing the challenge that providers have in getting people to collect the data; the inherent inequities in how surveys may be presented to different patients; and the implications that social determinants of health may have on whether a person responds. There are some concepts that HSAG may not be able to delve into deeply but may want to find a way to acknowledge whether the measure is trying to address them.
- A TEP member noted that his team has noticed in reporting more standard quality measures, there is a lack of infrastructure for effectively understanding their performance along racial and ethnic lines. He feels that the best performing measure in striving towards equity is one that specifically calls out reporting on performance along racial and ethnic lines. He suggested a scoring mechanism that gives the highest score to measures that call out some context-driven racial and ethnic reporting capabilities and perhaps a lower-tier score for measures that, while not reporting along specific racial and ethnic categories, may align with a condition or gap known to disproportionately affect racial and ethnic minorities. Examples would be hypertension or kidney disease.
    - HSAG Response: These are great thoughts for consideration. The team has been considering potentially investigating whether it would be feasible to give higher scores to stratified measures. It is somewhat of a challenge, in that not all measures are disparately sensitive and in terms of determining how the measures would be treated. The TEP member's idea about measures known to have very different outcomes among different populations is a good one. We would have to think through whether this can be scored or is more of a classification to bring to CMS' attention.
- A TEP member commented that she is impressed with how far the QMI has come and the robustness of the variables. She added that defining a high priority at a point in time offers flexibility and accommodates a lot of what she has heard in today's discussion. Indicating that many of the measures assessed by the project team are not publicly reported or have information that is not easy to access, she suggested the team consider the accessibility of measure data and availability of actionable information for the provider or patient. This is an important factor from a patient or consumer perspective.
    - HSAG Response: There are two cases for the QMI – the evaluation of existing measures in programs and evaluation of measures being considered for programs. At the consideration phase, it is not always known whether a measure will be publicly reported. However, for those measures already in use in an existing program, whether a measure's information is available to the patient could be considered for measure classification by the QMI.

Post-Meeting TEP comments

- A TEP member emailed the following comments after the meeting:
    - The team should continue to seek ways to assess whether measures are valuable to patients, families, and consumers. Usefulness to those receiving care is important for determining whether a measure is meaningful and usable; for instance, structural measures are very useful to the public but sometimes are given short shrift.
    - There is a growing concern related to data collection that provider burden is being shifted onto patients in the form of patient burden. Questionnaires that must be completed by patients prior to

a visit are growing exponentially, with duplication, reminders, and then texts within an hour of leaving the office to evaluate the visit.

- It is wrong to think that making performance data on highly technical quality indicators publicly available to consumers, e.g., through a website, would be helpful to most people. This idea does not reflect the body of research on the kinds of information that consumers find useful and actually rely on when making health care decisions or choosing providers. She advocates for taking this into account as the QMI is refined and improved.

Next Steps

Ms. Lockwood noted the 2022 CMS MDP QMI TEP nomination period is open until January 13, 2022. She encouraged members to reapply for the new TEP that will be in service from spring 2022 to spring 2024. She closed the meeting by thanking TEP members for their input and service.