Partnership for Quality Measurement
Powered by Battelle

Endorsement and Maintenance (E&M) Guidebook

October 2023

Prepared by:

Battelle
505 King Avenue, Columbus, Ohio 43201
Revised May 2023

The analyses upon which this publication is based were performed under Contract Number 75FCMC23C0010, entitled, "National Consensus Development and Strategic Planning for Health Care Quality Measurement," sponsored by the Department of Health and Human Services, Centers for Medicare & Medicaid Services.

Table of Contents

List of Tables

Page

List of Figures

Page

Introduction

The *Endorsement & Maintenance Guidebook (E&M Guidebook)* is a resource for measure stewards, measure developers, and organizations submitting measures to Battelle for the Partnership for Quality Measurement (PQM)™ for endorsement review. This Guidebook provides information about the various steps of the endorsement and maintenance (E&M) process, including each phased review, possible endorsement decision outcomes, the appeals process, E&M policies and procedures, and the E&M committee structure.

The E&M Guidebook is organized to provide an overview of E&M goals, priorities, and resources; to guide measure developers and stewards through the six steps of the E&M process; and to provide key considerations for submitting measures to Battelle. The E&M Guidebook aims to do the following:

- Explain the measure submission and evaluation processes
- Describe the expectations for measure developers and stewards as participants in the process
- Serve as the main resource for E&M-related processes and policies.

The E&M Guidebook will be updated on a timely basis to maintain a current reference to assist measure developers and stewards in navigating the E&M process.

Who We Are

Battelle is the world's largest independent, nonprofit, applied science and technology organization, with the objective of using science for the benefit of mankind. As a 501(c)(3) charitable trust, we are committed to translating scientific discovery and technology advances into societal benefits.

For over 20 years, we have been a leader in the science of health care quality measurement and improvement. Battelle is highly experienced in independent systematic evidence-based reviews of clinical quality measures (CQMs).

Battelle is a certified consensus-based entity (CBE) under the Centers for Medicare & Medicaid Services' (CMS) Qualified Entity (QE) Program developed to implement Section 10332 of the Affordable Care Act and the "Medicare Program; Availability of Medicare Data for Performance Measurement" Final Rule [CMS-5061-F]. As a certified CBE, Battelle meets the criteria of an independent CBE as mandated in federal statutes (SSA Section 1890 and 1890A).

To facilitate the execution of CBE tasks, we have formed the Partnership for Quality Measurement™ (PQM), which is comprised of all interested parties (formerly referred to as multi-stakeholder groups), including but not limited to health care providers (e.g., clinicians, health plans, health systems), patients and caregivers, measure experts (e.g., developers, stewards, researchers), and health information technology specialists. Battelle's transparent, streamlined approach to consensus-building facilitates informed and thoughtful endorsement reviews of quality measures. Membership in PQM is free, and to serve on an E&M committee, individuals must be members of PQM.

Battelle's Portfolio of CBE Measures

Battelle organizes measures for E&M by five (5) project topical areas, each having an evaluation committee, which oversees the portfolio of measures for the topic (Table 1). A project consists of measures submitted by measure developers/stewards and grouped by similar topic.

Table 1. Project Topical Areas

Project Title	Areas Covered	Example Measures *
Primary Prevention	Education, prevention, and screening related to health status and/or health risk	• CBE #0028 Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention • CB3 #2372 Breast Cancer Screening • CBE #3620 Adult Immunization Status

Project Title	Areas Covered	Example Measures*
Initial Recognition and Management	Recognition and timely diagnosis of conditions, including diagnostic accuracy, monitoring of early signs and symptoms of disease/condition	• CBE #0058 Avoidance of Antibiotic Treatment in Adults With Acute Bronchitis (AAB) • CBE #3671 Inappropriate diagnosis of community-acquired pneumonia (CAP) in hospitalized medical patients • CBE #3663e Excessive Radiation Dose or Inadequate Image Quality for Diagnostic Computed Tomography (CT) in Adults (Clinician Level)
Management of Acute Events, Chronic Disease, Surgery, Behavioral Health	Treatment of acute events, management of chronic disease, including structural or functional changes related to chronic disease, surgery and related outcomes	• CBE #0711 Depression Remission at Six Months • CBE #0729 Optimal Diabetes Care • CBE #3025 Ambulatory Breast Procedure Surgical Site Infection (SSI) Outcome Measure
Advanced Illness and Post-Acute Care	Advanced illness and/or end-stage disease management, palliative and hospice care, post-acute care, and home care	• CBE #0384e Oncology: Medical and Radiation - Pain Intensity Quantified • CBE #2651 Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Hospice Survey • CBE #2978 Hemodialysis Vascular Access: Long-term Catheter Rate
Cost and Efficiency	Total health care spending for a health care service or group of services associated with a specified patient population, time period, and/or unit of clinical accountability	• CBE #2158 Medicare Spending Per Beneficiary (MSPB) - Hospital • CBE #3575 Total Per Capita Cost (TPCC) • CBE #3561 Medicare Spending Per Beneficiary: Post Acute Care Measure for Inpatient Rehabilitation Facilities

** CBE # is the former National Quality Forum identification (ID) number. Battelle will continue with this numbering system.*

The categorization of endorsed measures from the 14 project areas under the prior CBE, the National Quality Forum (NQF), to the five new projects listed in Table 1 is not one-to-one. Rather, current and future endorsed measures are assigned to one of the five new projects based on the where in the patient's journey through health care the measure has the most relevance. Descriptions for the five new projects are listed below as well as on the PQM website:

- **Primary Prevention:** Primary prevention seeks to prevent disease or injury before it occurs. This can be achieved by altering unhealthy or unsafe behaviors that can lead to disease or injury and increasing resistance to disease should exposure occur.

This committee considers measures that address preventive care, education and counseling, health promotion, and screening measures. Subtopics may include health-related lifestyle behaviors, observable risk factors (e.g., age, obesity), and vaccination. Measures that address mental health screenings (e.g., depression, other psychiatric conditions), alcohol and drug use, comorbidities, and suicide risk screening are also included.

- **Initial Recognition and Management:** Timely recognition and diagnosis of certain conditions can lead to the use of effective treatments that work best at early stages of illness and overall improved management of disease.

  This committee considers measures that address early signs and symptoms of conditions and diseases, initial recognition and management of a change in a person's health status, and the experience, patient safety, and/or harm related to detection, diagnosis, or recognition of a condition. Measures in this topic area may include perinatal health, patient satisfaction with care, appropriate use of diagnostic tools, timeliness of care, care coordination, and monitoring.

- **Management of Acute Events, Chronic Disease, Surgery, Behavioral Health:** Disease management focuses on improving quality of life for individuals with acute and chronic conditions by minimizing or preventing the disease effects through quality care.

  This committee considers measures that address the management of acute or chronic disease, management of behavioral health conditions, pre-operative care, post-operative care, complications, functional status and impairments, patient safety, never events (i.e., medical errors should never occur), morbidity, and mortality. Measures in this portfolio may also focus on improving health care system care coordination.

- **Advanced Illness and Post-Acute Care:** With its focus on improving quality of life, advanced illness and/or post-acute care is distinct from care intended to cure an illness or condition, although it can be delivered concurrently with curative therapies. Advanced illness care focuses on medical care for people living with a serious illness. It aims to address the symptoms or stress related to the illness and improve the overall experience and/or quality of life for patients, caregivers, and family members. Advanced illness care may also include palliative and hospice care, which includes comprehensive care that addresses medical, emotional, spiritual, and social needs during the last stages of a person's terminal illness.

  Post-acute care includes medical or supportive care provided to individuals after leaving an acute care setting, such as a hospital, but they are not ready to return home. This includes rehabilitation and even palliative care services.

  This committee considers measures that address conditions including cancer and end-stage disease (e.g., end-stage renal disease). Subtopics may include advanced illness management, the assessment of physical, emotional, social, psychological, and spiritual aspects of care, access to and timeliness of care, patient and family experience with

care, patient and family engagement, care planning, hospice care, palliative and end-of-life care, chemotherapy, rescue, and intensive care unit (ICU) care.

- **Cost and Efficiency:** As health care expenditures continue to grow, it is crucial to understand how resources are utilized to maximize quality in the health care system. Health care cost measurement continues to be a critical component in assessing the United States (U.S.) health care system. Measures in the Cost and Efficiency portfolio are essential to evaluate the efficiency of care (i.e., higher quality, lower cost) and improve value through changes in practice. Improving U.S. health system efficiency can simultaneously reduce cost growth and improve the quality of care provided.

  This committee considers measures focusing on total health care spending for a health care service or group of services associated with a specified patient population, time period, and/or unit of clinical accountability. Measures in this topic area include broadly focused measures, such as per capita measures, which address total health care spending or resource use per person, to those with a narrower focus, such as measures dealing with the health care spending or resource use of an individual procedure (e.g., a hip replacement).

Submission Tool and Repository

Key information about measure submissions, including endorsement status, is available via the Submission Tool and Repository (STAR). STAR is an online platform where developers/stewards can submit measures and any interested party may view measure information, including the endorsement status, in the searchable repository database. The database is updated regularly as new and maintenance measures are submitted to Battelle for PQM endorsement review.

The measure submission function is available as of the Fall 2023 cycle. Measure developers/stewards must first create an account by going to the "Submit a Measure" page on the PQM website. Each cycle has a designated Intent to Submit deadline, where measure developers/stewards must submit key information (e.g., measure title, type, description, specifications) about the measure (*see Intent to Submit for more details*). One month after the Intent to Submit deadline (Table 2), measure developers/stewards submit the full measure information by the respective Full Measure Submission deadline (*see Full Measure Submission for more details*).

Table 2. Intent to Submit and Full Measure Submission Deadlines by Cycle

E&M Cycle	Intent to Submit*	Full Measure Submission *
Fall	October 1	November 1
Spring	April 1	May 1

** Deadlines are set at 11:59 p.m. (ET) of the day indicated. If the deadline ends on a weekend or holiday, the deadline will be the next immediate business day.*

To review the measure submission items and questions within each form, Microsoft Word templates of the Intent to Submit and Full Measure Submission forms are available on the PQM website and linked below:

- Intent to Submit Form (Version 1.0 | 2023)
- Full Measure Submission Form (Version 1.0 | 2023)

Endorsement and Review Process

Overview of and Enhancements to the Endorsement Process

The E&M process ensures measures submitted for endorsement are evidence-based, scientifically sound, and both safe and effective, meaning use of the measure will increase the likelihood of desired health outcomes; will not increase the likelihood of unintended, adverse health outcomes; and is consistent with current professional knowledge.

Our novel E&M process builds from the prior CBE processes and enables E&M decision-making in as few as six (6) months (from the Intent to Submit deadline until the end of the project [i.e., through the end of appeal proceedings]). Under this new process, measures reach their endpoint when an endorsement decision is rendered. This occurs when the E&M committees reach a final endorsement decision (Table 3).

Table 3. Endorsement Decision Outcomes

Decision Outcome	Description	Maintenance Expectations
Endorsed	**Applies to new and maintenance measures.** There is 75% or greater agreement for endorsement by the E&M committee.	Measures undergo maintenance of endorsement reviews every 5 years with an annual update review at 3 years (*see Annual Update for more details*).± Developers/stewards may request an extension of

Decision Outcome	Description	Maintenance Expectations
		up to 1 year (2 consecutive cycles).
Endorsed with Conditions*	**Applies to new and maintenance measures**. There is 75% or greater agreement that the measure can be endorsed as it meets the criteria, but there are recommendations/areas committee reviewers would like to see when the measure comes back for maintenance. If these recommendations are not addressed, then a rationale from the developer/steward should be provided for consideration by the E&M committee review.	Measures undergo maintenance of endorsement reviews every 5 years with an annual update at 3 years, unless the condition requires the measure to be reviewed earlier (*see Annual Update for more details*). The E&M committee evaluates whether conditions have been met, in addition to all other maintenance endorsement minimum requirements.
Not Endorsed°	**Applies to new measures only**. There is 75% or greater agreement to not endorse the measure by the E&M committee.	None
Endorsement Removed°	**Applies to maintenance measures only.** Either: • There is 75% or greater agreement for endorsement removal by the E&M committee; or • A measure steward retires a measure (i.e., no longer pursues endorsement); or • A measure steward never submits a measure for maintenance and there is no response from the steward after targeted outreach; or • There is no longer a meaningful gap in care, or the measure has plateaued (i.e., no significant change in measure results for accountable entities over time).	None

±Maintenance measures may be up for endorsement review earlier if an emergency/off-cycle review is needed (see Emergency/Off-Cycle Reviews for more details).

**Conditions are determined by the E&M committee, with the consideration of what is feasible and appropriate for the developer/steward to execute by the time of maintenance endorsement review.*

°Measures that fail to reach the 75% consensus threshold are not endorsed.

The "Endorsed with Conditions" category serves as a means of endorsing a measure, but with conditions set by the committee. These conditions should take into consideration what is feasible and appropriate for the developer/steward to execute by the time of maintenance endorsement review. The types of conditions that may be placed on a measure include but are not limited to:

- Conducting/providing additional testing across a larger population, accountable entity-level, and/or different level of analysis. This condition will mainly apply to new measures, especially those requiring more prospective testing (e.g., electronic clinical quality measures and patient-reported outcome performance measures).
- Expanding the measure use beyond quality improvement and into an accountability application (e.g., accreditation programs, pay-for-reporting programs, pay-for-performance programs, public reporting programs).

Battelle has identified several non-negotiable areas, meaning if a measure meets one or more of the following criteria, the measure cannot be endorsed, even with conditions:

- Lack of or unclear business case for the measure.
- Lack of evidence supporting the business case.
- Poor feasibility for the measure to be implemented due to challenges with data availability or missingness and/or due to substantial proprietary mechanisms prohibiting a measure's potential use.
- Inappropriate methodology, calculations, formulas, or testing approach used to demonstrate reliability or validity. This includes not testing the measure as specified.
- Specifications, testing approach, results, or data descriptions are insufficient for the committee to apply the PQM Measure Evaluation Rubric (*Appendix D*).

If a measure with an "Endorsed with Conditions" designation is evaluated for maintenance, but it has not met the prior conditions, then the committee may choose to remove endorsement, unless it agrees with any rationale provided by the developer/steward.

To achieve a 6-month E&M process (Figure 1) while maintaining high standards for transparency and rigor, we have enacted several key enhancements: 1) leveraging the Scientific Methods Panel (SMP) to advance measure science, 2) retiring the Consensus Standards Approval Committee (CSAC), 3) establishing a more robust and transparent appeals process, 4) leveraging the Novel Hybrid Delphi and Nominal Groups (NHDNG) technique[1], 5) reducing the number of E&M committees, and 6) conducting the pre-evaluation public commenting concurrent with staff assessments.

1. **Leveraging SMP to Advance Measure Science:** While each measure will continue to be reviewed for scientific rigor, the expertise of the SMP will be utilized across the entirety of the E&M process rather than on a measure-by-measure basis. While this does assist in reducing the overall endorsement timeline, more importantly it allows the SMP to provide guidance that enhances all measures by focusing on the most difficult methodological challenges faced by measure developers.

[1] Davies S, Romano PS, Schmidt EM, Schultz E, Geppert JJ, McDonald KM. Assessment of a novel hybrid Delphi and nominal groups technique to evaluate quality indicators. Health Services Research. 2011 Dec;46(6pt1):2005-18. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6773.2011.01297.x

2. **Retiring the Consensus Standards Approval Committee:** To empower E&M committees and to increase patient and consumer voices in the process, the CSAC review has been retired. Through PQM, endorsement decisions are made by E&M committees and are considered final unless an appeal is requested within three (3) weeks of the decision being posted to the PQM website. While this does assist in reducing the overall endorsement timeline, with this change, we have also increased the engagement of the patient community (i.e., patients, caregivers, advocates) and purchasers of health care (e.g., state, federal, private) in the E&M process. In particular, at least three members from the patient community are seated on each committee, increasing the reach to all interested parties, including patients and purchasers through our partnerships with IHI (Institute for Healthcare Improvement), Rainmakers, and PQM membership overall.
3. **Establishing a More Robust and Transparent Appeals Process:** When an appeal is received, the E&M team conducts a preliminary review of the appeal to determine if a re-review is appropriate, at which point an ad-hoc Appeals Committee is convened to review the appeal. The Appeals Committee consists of the E&M team (i.e., technical leaders from PQM and our partners) and the co-chairs from each of the five (5) E&M committees. If additional perspectives are needed, we will recruit needed subject matter experts (SMEs) to support the Appeals Committee review. This structure ensures these meetings can be convened quickly and as needed, and the inclusion of E&M staff and committee chairs reduces the risk of duplicative or contradictory discussions (*see Appeals for more details*).
4. **Leveraging the Novel Hybrid Delphi and Nominal Groups Technique:** To promote consistency in measure evaluation reviews and to ensure focused, facilitated discussion that is inclusive of all interested party perspectives, Battelle uses a NHDNG technique to build consensus among E&M committee members and leverages experienced and trained facilitators. The NHDNG technique utilizes a multi-step process meant to increase engagement of all committee members and structure measure review facilitation to build consensus.
5. **Reducing the Number of E&M Committees:** Measures submitted to Battelle for review are evaluated by one of the five E&M project committees (Table 1). Therefore, the previous endorsement committees have been retired, except for Cost and Efficiency. Maintenance of fewer, more generalized committees enables a more equitable distribution of effort. These committees have a diversity of expertise to address a range of topic areas in a more flexible approach that maximizes engagement during each cycle. SMEs are recruited, as needed, to provide more specific clinical knowledge when called for by the measure under review (*see E&M Committee Composition, Roles, Responsibilities for more details*).
6. **Conducting the Pre-evaluation Public Commenting Concurrently with Staff Assessments:** For each measure evaluation cycle, public comment happens twice. The first occurs prior to the E&M committee endorsement meeting (i.e., pre-evaluation commenting), and the second occurs after an endorsement decision has been made by the committee (i.e., post-evaluation commenting). The first comment period has the lengthier public comment window of 30 days, which enables 1) public comment to happen concurrent with an internal E&M team review and assessment (formerly Preliminary Analysis); and 2)

the compilation and synthesis of comments received, which is integrated into the NHDNG process for review during the E&M committee meetings before members make their final endorsement decisions about the measure(s).

For the second public comment period, endorsement decisions made during committee meetings are shared via the public-facing PQM website for three (3) weeks, which represents the Appeals Period. During this period any interested party may request an appeal of any E&M committee endorsement decision rendered for that cycle.

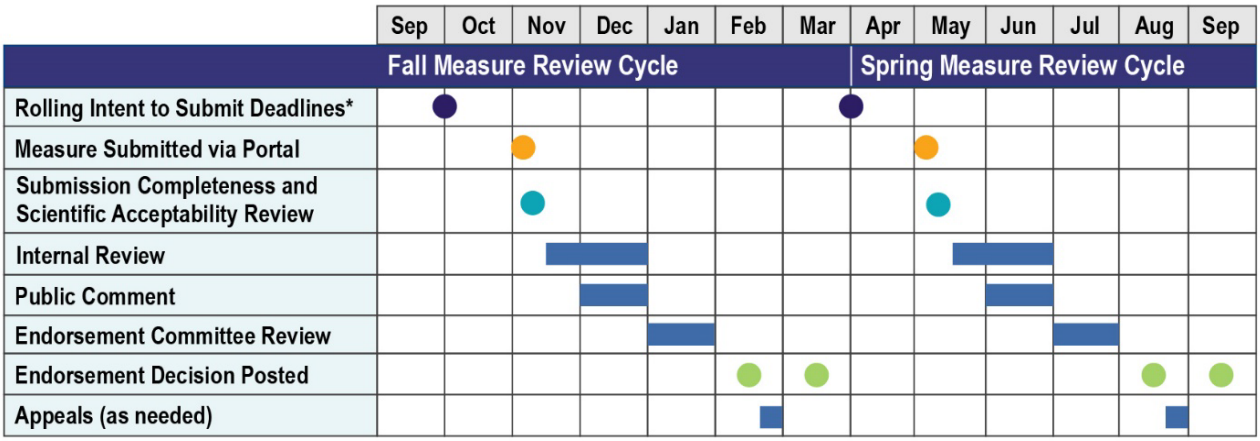


Figure 1. A 6-month endorsement review process.

Battelle is dedicated to evaluation of the E&M process and integrating interested party input across the cycles. Therefore, the process and the E&M Guidebook may evolve over time. Proposed changes to the E&M process or criteria undergo a formal public comment period; before any changes are implemented, educational resources (e.g., webinars, informational guides) are made available. Any changes in the E&M process or measure evaluation criteria will not be applied to any measure that is currently going through the E&M process.

Prior to the Fall 2023 cycle, we hosted an informational webinar on June 30, 2023, to review the enhancements of the E&M process, including an overview of the E&M Guidebook. We also announced a subsequent public comment period from July 1 to 30, 2023, to solicit feedback on various aspects of the E&M Guidebook and process overall. All comments were posted to the PQM website, and our responses have been provided in Appendix G, noting where changes have been made to the E&M Guidebook. This version of the E&M Guidebook applies to the E&M process starting with the Fall 2023 cycle.

This E&M Guidebook will be updated on a timely basis to maintain a current reference to assist measure developers and stewards in navigating the E&M process. Any major updates to this E&M process, policies, evaluation rubric, etc. will be shared for public comment prior to implementation of any changes.

E&M Committee Composition, Roles, Responsibilities

As a CBE, Battelle seats individuals from the PQM membership into committees to participate in the E&M process. E&M committees are composed of diverse members representing all facets of the health care system. There are up to five projects each cycle, each having a committee that evaluates, discusses, and assigns ratings and endorsement decisions for measures under endorsement review (see Table 1). Battelle ensures diversity of E&M committee membership through a formal nominations process (*see E&M Committee Nominations for more details*) to fill gaps in expertise and perspective needed for the E&M committee. This structure of membership organization enables use of the NHDNG technique, which maximizes member engagement and promotes consistent application of evaluation criteria.

Each project uses the Novel Hybrid Delphi (Advisory) and Nominal (Recommendations) Groups technique for measure endorsement reviews. Each E&M project committee is divided into an Advisory Group and a Recommendations Group (Figure 2).

- **Advisory (Delphi) Group**: Members in this group review and provide ratings and written comments on measures prior to the Recommendations Group endorsement meeting. These inputs ensure a larger number of voices contribute to the consensus-building process. The Advisory Group members also attend the Recommendations Group endorsement meeting to listen to the Recommendations Group discussions and to vote on measure endorsement decisions during the meeting.
- **Recommendations (Nominal) Group**: Members in this group also review and provide ratings and written comments on measures prior to the Recommendations Group endorsement meeting. Areas of disagreement (i.e., lack of consensus) identified from the initial measure ratings from both groups will inform the Recommendations Group discussions during the endorsement meeting (*see Endorsement Committee Review for more details*). Recommendations Group members will also vote on measure endorsement decisions during the meeting.

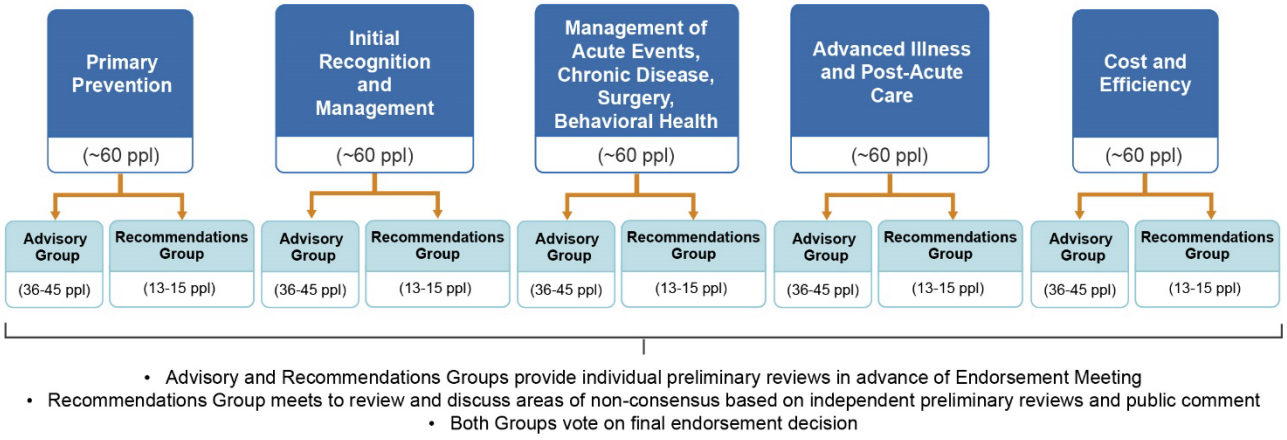


Figure 2. Recommendations and Advisory Group Structure

The Recommendations and Advisory Group members consist of interested parties from PQM membership, who evaluate, discuss, and rate the measures undergoing endorsement review.

Each E&M project committee (the Recommendations Group plus the Advisory Group) has two co-chairs, who participate in the Endorsement meeting and take part in the Recommendations Group discussions. When possible, we ensure at least one co-chair is from the patient community. The patient representative co-chair is responsible for engaging and supporting patient representatives on their respective committee. The other co-chair is responsible for ensuring the concerns and perspectives from the Advisory Group are considered by the Recommendations Group during the endorsement meeting. In addition, the co-chairs' responsibilities are to:

- Co-facilitate endorsement meetings, along with E&M project staff
- Work with E&M staff to achieve the goals of the project
- Assist E&M staff in anticipating questions and identifying additional information that may be useful to the committee
- Participate in the Recommendations Group as a full voting member for the entirety of their term
- Serve on the Appeals Committee.

To ensure representation of the population of interested parties, up to 60 members are seated on an E&M project committee through a formal nominations process (*see E&M Committee Nominations for more details*), which is conducted annually to fill gaps in expertise and roster categories (Table 4). To serve on an E&M committee, individuals must also be PQM members. We seat PQM members based on the expertise needed for the E&M project, ensuring adequate representation and perspectives across roster categories. Additionally, all newly seated committee members are invited to an annual E&M virtual orientation meeting, which provides an overview of the E&M process, committee roles and responsibilities, as well as a review of the PQM Measure Evaluation Rubric (*Appendix D*).

As needed, the membership of the Recommendations Group may be augmented with individuals with specialized expertise, which is determined after each cycle's Intent to Submit deadline. For example, if a health care cost measure for a specific disease state or condition is under review by the Cost and Efficiency committee, SMEs familiar with that disease state or condition are invited to the endorsement meeting to provide further context and relevance for the committee's consideration. These SMEs are non-voting participants and will only provide input on relevant measures.

If additional expertise is needed, we first identify if the needed expertise resides within one of the other E&M committees. If SME expertise is still absent from other E&M committees, we and our partner, IHI, recruit SMEs from our combined networks, and selection is vetted against our conflict of interest policy and with input from the respective committee co-chairs. If needed, we may establish a pool of SMEs across various clinical (e.g., nephrologists, primary care providers) and methodological (e.g., psychometricians) areas. This pool of SMEs will also have term limits and are subject to the same conflict of interest policy as committee members (*see Conflict of Interest Policy for more details*).

Table 4. Roster Categories and Target Number of Individuals

Roster Category	Advisory Group Targets*	Recommendations Group Targets*
Patients, families, caregivers, patient advocates	9	3
Clinicians, including physicians, nurses, pharmacists, physical therapists, etc.	6	2
Facilities/institutions including accountable care organizations, hospitals or hospital systems, and post-acute/long-term care facilities	6	2
Purchasers and plans (state, federal, and/or private)	6	2
Rural health experts	3	1
Health equity experts	3	1
Researchers in health services, alternative payment models, population health	6	2
Other interested parties (representatives of electronic health record [EHR] vendors, provider and facility associations, and experts in areas such as quality improvement/ implementation science, care coordination, patient safety, behavioral health, and national policy makers)	6	2
TOTAL	**45**	**15**

**Note: If Battelle does not fill the number of seats listed for a given roster category, Battelle will determine if remaining seats can be distributed to other roster categories, based on the expertise needed within the committee.*

Term of Appointment

Committee members are appointed to a 3-year term. However, for the first year of the new E&M process (i.e., Fall 2023 and Spring 2024 cycles), committee members were randomly assigned, within roster categories, to a 1-, 2-, or 3-year term. Those individuals who received a 1-year term were assigned to the Recommendations Group for their entire term. Members with 2- and 3-year terms were assigned to the Advisory Group. As the Recommendations Group members end their 1-year terms and roll off the committee, the 2-year term Advisory Group members move into the Recommendations Group, and the 3-year Advisory Group members remain within the Advisory Group but are the next members to serve on the Recommendations Group. Through yearly committee nominations, newly appointed committee members are assigned a 3-year term and initially seated to the Advisory Group. This approach provides a staggering of committee member term limits and facilitates a continuity of process and content knowledge within committees. Furthermore, this staggering approach ensures each member of the

Advisory Group will have the opportunity to serve on the Recommendations Group within their term. With respect to the co-chairs, for the first year, one co-chair was randomly assigned a 1-year term and the second co-chair to a 3-year term. Co-chairs always serve on the Recommendations Group throughout their terms.

We ensure no more than one-third (1/3) of members roll off the committee every year. Committee members who roll off the committee, including co-chairs, may reapply to serve on the committee during the call for nominations. There is no requirement that former committee members wait to reapply.

E&M Committee Nominations

We conduct a review of committee member appointments annually, which includes internal re-calibration of the membership, a call for nominations, and targeted outreach. A call for nominations is published on the PQM website and an announcement is sent out to all PQM members. Nominations are submitted via the PQM website. Self-nominations are welcome. Third-party nominations must indicate that the individual has been contacted and is willing to serve. Nominees must be PQM members (which is free), and they must complete an application form and a Personal/Organizational Disclosure of Interest (DOI) form (*Appendix B*). Before finalizing the appointments, a draft roster of nominees is published for public comment for transparency and for garnering input as to whether the E&M roster has the expertise needed for the given E&M project.

Nominees commit to participating in scheduled calls and meeting dates, providing timely responses to requests for feedback, and being available for ad-hoc meetings and conference calls. To be eligible for participation, nominees should (1) have relevant expertise and demonstrated experience related to the use of quality and efficiency measures and/or (2) belong to at least one of the following categories:

- Patients, caregivers, and patient advocates
- Clinicians, including physicians, nurses, pharmacists, physical therapists, etc.
- Facilities/institutions including accountable care organizations, hospitals or hospital systems, and post-acute/long-term care facilities
- Purchasers and plans (state, federal, and/or private)
- Rural health experts
- Health equity experts

- Researchers in health services financing, alternative payment models (e.g., bundled payment, shared savings, all-payer models, etc.), population health, or implementation science methodology
- Other interested parties (representatives of electronic health record [EHR] vendors, provider and facility associations, and experts in areas such as quality improvement/ implementation science, care coordination, patient safety, behavioral health, and national policy makers.

Committee members are responsible for notifying the E&M project team if they:

- Change employers or contact information;
- Are unable to attend a scheduled meeting; and/or
- Have a prolonged conflict emerge during their term that will interfere with meeting the obligations of E&M committee membership, in order to determine whether ongoing membership on the committee is warranted or if inactive status can be granted for a cycle.

Inactive Status and Early Termination

We understand plans and demands of our volunteer E&M committee members change. Therefore, members may need to move to inactive status for a given review cycle or end their terms early. E&M committee members with inactive status continue with their terms, but for the cycle of interest, they are not permitted to vote and are therefore not counted in the denominator when determining meeting quorum and voting thresholds. A committee member may be granted inactive status at any time before the endorsement meeting.

If a committee member has poor attendance or participation, as determined by not attending one or more endorsement meetings without advance notice and/or by not submitting independent reviews of measures for endorsement review (*see Independent E&M Committee Member Review and Assessment for more details*), we will contact the member and ask if he/she would like to resign. We reserve the right to remove any member from an E&M committee, including for reasons of persistent poor attendance or lack of participation.

Conflict of Interest Policy

As a Consensus Based Entity (CBE) on contract 75FCMC23C0010 with CMS, Battelle Memorial Institute (“Battelle”) convenes several committees of interested parties to provide input on (1) endorsement decisions on quality performance measures; (2) the selection of quality measures for a pre-rulemaking process, which is required by Social Security Act Sections 1890(b)(7) and 1890A; and (3) a measure removal process. This Conflict of Interest Policy (the “Policy”) is applicable to such committees to ensure each committee performs its functions in a manner free from bias and undue influence. All committee members must attest they will follow this policy and provide the requisite information necessary for Battelle to conduct a conflict of interest (COI) review.

The term “conflict of interest” means any financial or other interest that could actually or be perceived to (1) significantly impede your objectivity, or (2) create an unfair competitive advantage for you or an organization associated with you. Disclosure of a financial interest does not automatically mean a COI exists but may warrant further discussion and review.

As part of the E&M committee nomination process, each nominee completes a Personal/Organizational Disclosure of Interest form (*Appendix B; see E&M Committee Nominations above*). In addition, to complete the COI analysis, each member serving on a committee evaluates measures for endorsement and/or for providing recommendations for pre-rulemaking will be required to complete a Measure Disclosure of Interest Form for each measure, or batch of measures, assigned to that committee (*Appendix C*). This form will contain questions relevant to the specific measure(s) being reviewed. Battelle will provide the Measure Disclosure of Interest Form to committees at the start of each cycle. The form will contain questions regarding the member’s financial interests and business associations, which may present a perceived or actual COI.

The questions in the Measure Disclosure of Interest Form will focus on whether:

(1) You contributed directly and substantially to the development of a measure or measures being considered for endorsement or under consideration for selection or removal. For example,

- You worked on the measure as an employee of or consultant for the measure development organization
- You directly collaborated with the measure development organization to create or refine the measure.

(2) You or your spouse, domestic partner, or child could receive a direct financial benefit from a measure being recommended for selection, removal, or endorsement. For example,

- You own stock in a company that has a financial interest in the measure being endorsed or not endorsed.

(3) In the last five (5) years you have received an indirect financial benefit, i.e., not related to the measure under review, of $10,000 or more from a measure developer whose measure is under review, or an indirect financial benefit of $10,000 or more, in the aggregate, from an organization or individual that may benefit from a measure being endorsed or not endorsed or being considered for the selection or removal process. For example,

- You have received $20,000 in consulting fees from the measure developer in the last five (5) years for work unrelated to the measure being reviewed.

(4) You are currently employed by the measure developer and the developer has created the measure(s) under review, has created measure(s) in the topical area under review, or

has created measure(s) that compete with measure(s) created by another developer and under review.

By participating as a committee member, each member consents to public disclosure of general information about the member's financial or business interests, professional associations, and experiences that may be of interest to the public regarding COI. Members must also announce their organizational affiliation and any organizational conflicts of interest. Unless legally required to do so by an authoritative entity, such as CMS, specific financial information will not be provided to the public, but financial relationships may be subject to disclosure.

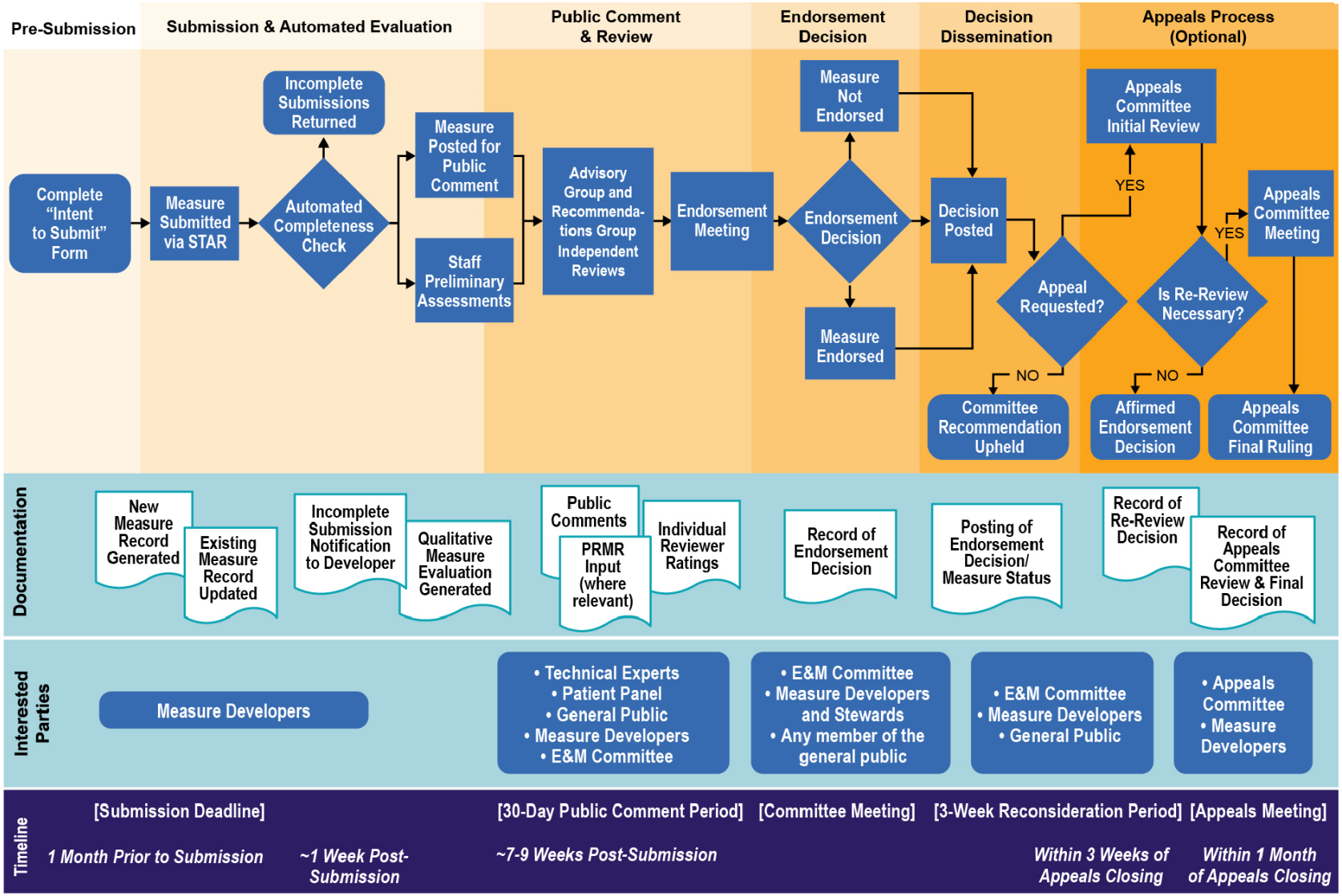
If you provide information that creates a perceived or actual COI, Battelle requires you recuse yourself from any voting regarding the applicable measure or measures, and in some instances, competing and related measures. However, you may still contribute to the discussion of the measure. Committee members who have conflicts with specific measures, as determined by the Measure Disclosure of Interest Form, must publicly recuse themselves from any voting associated with those measures.

Additionally, committee members must orally disclose relevant interests at a public committee meeting. The disclosure usually occurs at a committee's endorsement meeting. Senior Battelle staff will lead this disclosure and instruct committee members regarding information that should be disclosed. Following oral disclosure by committee members, Battelle staff will invite committee members to ask and respond to questions of each other or Battelle staff regarding any disclosures made by committee members.

Finally, all committee members have an ongoing duty to monitor for COI issues of themselves and fellow committee members and raise or disclose any issues either in a committee meeting, to the committee chair, the Battelle program team, or the Battelle legal department. Committee members should take a proactive approach and report any instances if a fellow committee member appears conflicted or is acting in a biased manner.

Submitting Measures to Battelle

The E&M process consists of series of stages, starting with Intent to Submit and cascading to the Appeals period (Figure 3). During each stage, we work closely with developers and stewards, committee members, and other interested parties to address questions regarding process and/or criteria. We also conduct the endorsement meetings and provide all relevant materials and documentation of the endorsement deliberations and decisions, including committee rationales. We inform all interested parties of the status of measures going through the process and welcome public comment on the measures and endorsement decisions throughout the review cycle. Lastly, all information pertaining to the E&M committee meetings, the measures being reviewed, and the E&M meetings themselves are made public.



*Measures submitted for maintenance follow the same process as new measures but have different criteria for endorsement at internal measure review & committee review stages.
PRMR: Pre-rulemaking measure recommendations

Figure 3. Process map of one, 6-month E&M cycle.

Intent to Submit

The Intent to Submit period is on a rolling basis. Measure developers and stewards submit key information about the measure via STAR at least one (1) month prior to the full measure submission deadline of the intended review cycle (Fall or Spring, year). For all measures (new and maintenance), stewards/developers must submit the following information during Intent to Submit:

- Measure Title
- Measure Description
- Measure Type (e.g., structure, process, outcome)
- Measure Specifications (e.g., numerator, denominator, exclusions, level of analysis, care setting)
- Intended Measure Review Cycle
- Contact information and affiliation
- Attestations for what is required by Full Measure Submission

Throughout and leading up to the intended measure review cycle, developers and stewards may request technical assistance, which we can provide (*see Technical Assistance for more details*).

Full Measure Submission

Completeness Checks

After one (1) month from Intent to Submit, developers/stewards must submit all the measure information via the online measure submission function of the STAR. Requirements for initial and maintenance measure endorsement are indicated as, "*[For initial endorsement]*" or "*[For maintenance]*," within each domain of PQM Measure Evaluation Rubric (*Appendix D*). If neither distinctions are listed for a rubric requirement, then it applies to both initial and maintenance endorsement.

We conduct completeness checks (see *Measure Submission Completeness Checklist* below) to determine if all required responses and measure information have been submitted. We notify measure developers/stewards of any issues identified and request developers/stewards address the completeness check feedback by the requested deadline, which is no less than two (2) business-days from receipt of the completeness check feedback. Measures that pass the completeness check review are posted for a 30-day public comment period, while simultaneously undergoing an internal measure review by E&M staff (see *E&M Team Preliminary Assessment* for more details).

Submission of Electronic Clinical Quality Measures (eCQMs)

The following clarifications are specific to eCQMs:

- A new eCQM version of an endorsed measure is not considered an endorsed measure until it has been specifically evaluated and endorsed by Battelle. An eCQM should be submitted as a separate measure even if the same or a similar measure exists.
- Measure specifications should use the latest accepted versions of the following industry eCQM technical specifications: Health Quality Measure Format (HQMF), Quality Data Model (QDM), and Clinical Quality Language (CQL). Use of the CMS Measure Authoring Tool (MAT) ensures the measure uses these technical specifications; however, the MAT is not required to produce HQMF.
- eCQM developers must use value sets that are published through the National Library of Medicine's Value Set Authority Center (VSAC). This helps reduce implementation issues related to value sets and code system validation and encourages the use of harmonized value sets. If an eCQM does not have a published value set, then the measure developer must look to see if there is a published value set that aligns with the proposed value set within its measure. If such a published value set does not exist, then the measure developer must demonstrate the value set is in draft form and is awaiting publication to VSAC.
- Testing within electronic health record (EHR) systems from at least two (2) EHR vendors. Beyond this minimum requirement, developers/stewards should test on the number of her systems they deem appropriate. Submission requirements for eCQMs also include a feasibility assessment, using the eCQM Feasibility Scorecard. This assessment identifies data element feasibility issues. Simulated data set results allow assessment of each branch of the measure logic to ensure the logic can be processed technically by other eCQM-capable reporting tools.
- Empirical demonstration of data element reliability is required for unstructured data fields and data element validation is required for all eCQMs. If the testing is focused on validating the electronic data elements, developers/stewards should analyze agreement between the electronic data obtained using the eCQM specifications and those obtained through abstraction of the entire electronic record (not just the fields used to obtain the electronic data). Developers/stewards should use statistical analyses, such as sensitivity and specificity, positive predictive value, and negative predictive value. This type of validity testing also satisfies the requirement for reliability testing. If data element testing is not possible, justification is required and must be accepted by the E&M committee. Face validity alone will not be sufficient.

Measure Submission Completeness Checklist

Developers/stewards are also encouraged to follow the checklist below to ensure the measure submission is complete and responsive prior to E&M committee consideration. The E&M team will review measure submissions for completeness. If issues have been identified, the E&M

team will notify developers/stewards within one week after the measure has been submitted. Developers/stewards will be asked to address the completeness check feedback by the requested deadline, which will be no less than two (2) business days from receipt of the feedback. If developers/stewards are unable to do so, then the measure may need to be pulled from consideration and resubmitted at a future cycle.

- ☐ Quality Measure Developer and Steward Agreement (QMDSA) completed and signed (*Appendix A*).
- ☐ Health conditions (i.e., disease states) of the measure are indicated.
- ☐ Adequate responses received for all relevant and required fields within the measure submission form.
- ☐ Testing is conducted for the data source(s) and level(s) of analysis for which the measure is specified; information for data source and level of analysis is consistent across the specifications and testing items.
- ☐ Attachments, including electronic clinical quality measure (eCQM) specifications, Feasibility Scorecard, and data dictionary/code lists, if applicable and appropriate.
- ☐ All URLs are active and accurate.
- ☐ Evidence of a search for potential related/competing CBE measures and either a plan for harmonization or a rationale is provided for identified measures.
- ☐ All measure submission information, including attachments, is 508-compliant (*see Appendix E for more details*)
- ☐ Paired measures are submitted on separate forms.
- ☐ ICD-10 (International Classification of Diseases) codes are used and included, if applicable.

E&M Team Preliminary Assessment

We review each measure submission using the PQM Measure Evaluation Rubric (*Appendix D*). Measures are evaluated on five domains (Importance, Feasibility, Scientific Acceptability [i.e., Reliability and Validity], Equity, and Use and Usability. For each domain, we indicate if a measure domain has been “Met”, “Not Met but Addressable”, or “Not Met”, based on specific evaluation considerations for each area. These preliminary assessments summarize key points of the submission as they pertain to the PQM Measure Evaluation Rubric, and when appropriate, provide additional context or interpretation for certain aspects of the submission (e.g., verifying a testing methodology is appropriate). The preliminary assessment ratings are not binding, but instead, are meant to serve as input for committee discussion.

We share these preliminary assessments (formerly called “Preliminary Analysis”) with developers/stewards for a factual review prior to sharing with the E&M committee.

Developers/stewards are asked to conduct a factual review by the requested deadline, which is no less than two (2) business days from receipt of the preliminary assessment. This factual review is to ensure the preliminary assessments include accurate results from the measure submission. For example, when summarizing the testing results of a measure, have we accurately reflected the testing results? This factual review *is not* intended to provide an opportunity for developers/stewards to disagree with the preliminary measure ratings.

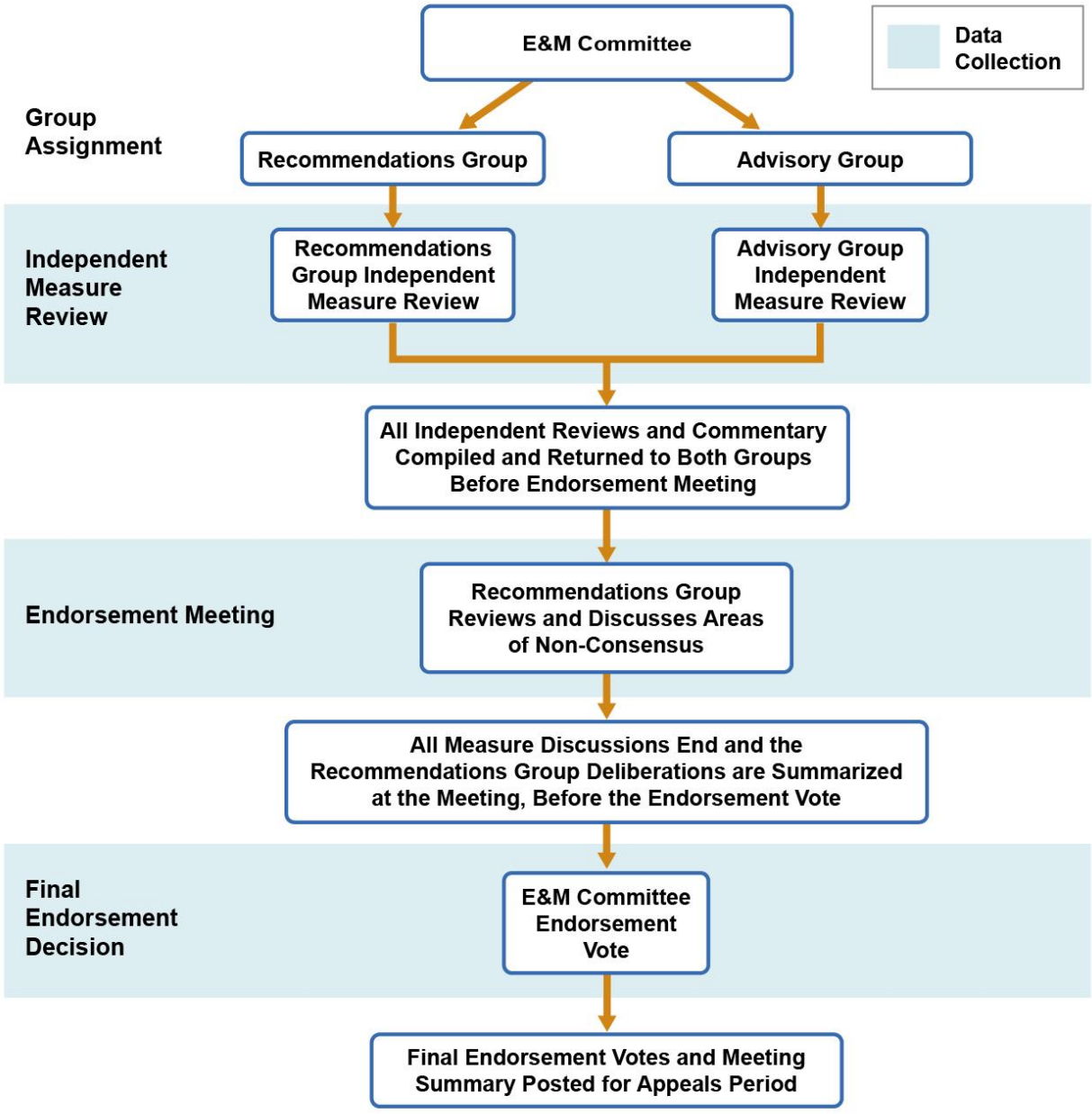


Figure 4. Advisory and Recommendations Group Measure Review Process

Independent E&M Committee Member Review and Assessment

At least three (3) weeks prior to an E&M committee endorsement meeting (referred to as "Endorsement Meeting" in Figure 4), the Recommendations Group and the Advisory Group of each E&M committee receive the full measure submission details for each measure up for review, including all attachments, the PQM Measure Evaluation Rubric, a link to the public comments received for the measure(s), and the E&M team preliminary assessments.

Members of both groups are asked to review each measure, independently, against the PQM Measure Evaluation Rubric (referred to as "Independent Measure Review" in Figure 4). Committee members assign a rating of "Met", "Not Met but Addressable", or "Not Met" for each domain of the PQM Measure Evaluation Rubric. In addition, committee members provide associated rationale for each domain rating, which is based on the rating criteria listed for each domain (*see Appendix D for more details*). We aggregate and summarize the results and distribute them back to the committee, and to the respective measure developers and/or stewards, for review within one (1) week of the endorsement meeting.

These independent committee member ratings are compiled and used by Battelle facilitators and committee co-chairs to guide committee discussions (*see Endorsement Meeting section for more details*).

Public Commenting

For each measure evaluation cycle, there are three (3) public comment opportunities, prior to the endorsement meeting, during the endorsement meeting, and after the endorsement meeting (i.e., the Appeals Period).

Public Comment Prior to the Endorsement Meeting

Once a measure submission passes the completeness check step, the full measure submission details, including all attachments, are posted to the PQM website for a 30-day public comment period. This first public comment period occurs prior to the endorsement meeting and concurrently with the development of the E&M team's preliminary assessments. The intent of this 30-day comment period is to solicit both supportive and non-supportive comments with respect to the measure(s) under endorsement review. Any interested party may submit a comment on any of the measures up for endorsement review for a given cycle (e.g., Fall or Spring). All public comments received during this 30-day period are posted to the PQM website for full transparency. We share all comments received with each respective E&M committee for consideration during its independent committee member review. We also summarize the comments received during the endorsement meeting.

Developers and/or stewards are not asked to respond to the pre-endorsement meeting comments in advance of the endorsement meeting. However, developers/stewards may do so at their discretion as part of their opening remarks and/or at designated points during the endorsement meeting.

Public Comment During the Endorsement Meeting

The second public comment opportunity occurs during the committee endorsement meeting. After the committee has ceased its discussion of a measure and before the committee votes on an endorsement decision for a measure, the E&M team will open the floor for public comment. The intent of this public comment opportunity is for any interested party to verbally submit a comment related to the committee measure review deliberations. After being recognized, commenters should clearly state their name and affiliation (if applicable) before stating their comment. Lastly, commenters will have no more than five (5) minutes to provide their comment for the committee's consideration.

The committee is not tasked with addressing or adjudicating these comments. Rather, the committee listens to comment, noting any support for or concerns with the measure(s) and/or committee deliberations. If needed, the committee may have a final round of discussion, based on the verbal comments received, before moving to an endorsement vote.

Public Comment After the Endorsement Meeting (i.e., the Appeals Period)

The third public comment opportunity occurs after the endorsement meeting, when all E&M committee endorsement decisions are shared publicly via the PQM website for three (3) weeks. The intent of this public comment period is to provide an opportunity for any interested party to submit a comment requesting for an appeal of any committee endorsement decision. Referred to within this E&M Guidebook as the Appeals Period, any appeal received must be eligible before being considered by the Appeals Committee (*see the Appeals section for more details*).

Endorsement Meeting

Novel Hybrid Delphi and Nominal Groups Technique

After the independent committee member reviews have been aggregated and distributed back to the respective E&M committees, the committee's Recommendations Group is convened for the endorsement meeting. One endorsement meeting, spanning 1—2 days, is held per project per cycle, and all meetings are held virtually and open to the public. The committee's Advisory Group members also join the endorsement meeting to listen to the Recommendations Group discussions and to provide an overall endorsement vote. Measure developers and/or stewards are also invited to the endorsement meeting to introduce their measure(s), to provide further context and rationale for their measure(s), and to answer questions posed by the committee during designated times.

During the meeting, the E&M team and committee co-chairs review the results from the committee independent reviews, E&M team preliminary assessments, and public comments received. Through facilitated discussion, the Recommendations Group focuses on aspects of the measure which raised concern, as identified through the independent committee reviews, public comments, and/or E&M team preliminary assessments. After the discussions cease for a measure, the E&M committee co-chairs summarize the deliberations of the Recommendations Group before moving to an endorsement vote. Within this summary, co-chairs draw attention to the issue(s) discussed, noting the recommended endorsement conditions, if any, and clearly

summarizing the committee rationales for supporting and not supporting the measure. These rationales are based on whether PQM Measure Evaluation Rubric domains are “Met”, “Not Met but Addressable”, or “Not Met.”

The committee does not vote on each domain of the PQM Measure Evaluation Rubric. Rather, both the Advisory Group and Recommendations Group members only vote on the endorsement decision (Table 3). The committee provides its rationale for each endorsement decision, noting any deficiencies in the submission or the measure specifications or the failure to identify deficiencies in the submission. For example, if committee members determine the Scientific Acceptability is "Not Met" due to concerns with the reliability testing results being low, then this rationale, plus any other issues identified, is clearly stated with the endorsement vote. If the committee does not reach consensus on an endorsement vote, then the measure is not endorsed.

Consensus is determined to be 75% or greater agreement among members. Consistent with the goal of adding rigor to all aspects of the consensus development process, Battelle established the 75% threshold of consensus from an evidence-based consensus measure. Analogous to inter-rater reliability statistics, the consensus measure allows the E&M team to assess the degree of disagreement (or lack of consensus) amongst the independent committee reviews and the committee endorsement votes. The consensus measure is an index of agreement, where the closer to one (1), the more agreement, or consensus, there is. From the table in Appendix F, when the consensus measure is 0.95 or greater, the corresponding threshold of consensus is 75%. This approach is advantageous compared to other metrics based on variance, in that it takes into consideration the different sizes of the voting groups and different rating options (*see Appendix F for more details*). Figure 5 below depicts how endorsement decisions are reached based on the 75% consensus threshold.

Endorse (A)	Endorse with Conditions (B)	Do Not Endorse (C)	Consensus Voting Status
75% or More	0%	Less than 25%	A
75% or More		Less than 25%	B
Less than 25%		**75% or More**	C
26% to 74%		26% to 74%	No consensus

Figure 5. Consensus Voting for Final Endorsement Decisions

After the meeting, we post a meeting summary, including all endorsement decisions and associated rationales, to the PQM website for a three (3)-week Appeals Period. We coordinate all communications with the Recommendations and Advisory Groups and coordinate all meetings. After the close of the Appeals Period, we post a final technical report to the PQM website for each E&M project convened during an endorsement cycle (*see Final Technical Report for more details*).

A crucial aspect of a successful consensus-based process is effective and organized meeting facilitation to ensure discussions remain productive, within scope, and inclusive of all voices. At E&M committee meetings, Battelle facilitators confirm quorum and engage committee members in robust discussion to build consensus recommendations about each measure under review. The NHDNG is a comprehensive, adaptable tool that is employed to build consensus among E&M committee members and leverage experienced and trained facilitators. Further, the NHDNG technique is a hybrid technique, utilizing a multi-step process meant to increase engagement of all committee members and structure facilitation by using standard measure evaluation criteria and practices.

The **consensus-based process** ensures:

- Productive discussions
- Discussions within scope
- Inclusion of all voices
- Increased engagement
- Efficient information exchange

This structured approach allows for efficient information exchange among E&M committee members, which is particularly important when members offer unique points of view. The use of independent committee reviews anchors opinions based on each committee member's knowledge and limits the likelihood that a vocal few impart too much bias on the results. Furthermore, this approach ensures all members have access to the same information prior to final evaluation.

Quorum

Having a quorum for meeting attendance and for voting is critical to ensuring the Recommendations Group discussions and the E&M committee vote are robust and reflective of all perspective represented on the E&M committee. Meeting quorum requires at least 60% of the Recommendations Group members only to be present during roll call at the beginning of the meeting. Meeting quorum is not established for the Advisory Group members, as they do not discuss the measures under review during the endorsement meeting. However, attendance of both groups is taken into consideration when establishing whether voting quorum is met.

We monitor attendance during the meeting to ensure both meeting quorum and voting quorum are maintained. If there is less than 60% attendance, then the Recommendations Group will not discuss the measures and a back-up meeting will be held. If meeting quorum is lost during the meeting, the measure evaluation discussions will cease and an alternative meeting will be held to complete the measure review.

Voting quorum is at least 80% of all active committee members (the Recommendations Group plus the Advisory Group), who have not been recused (*see Conflict of Interest Policy for more details*). If the voting quorum is not met at committee roll call but meeting quorum is achieved, the Recommendations Group proceeds to discuss the measures, but endorsement voting will not occur during the meeting. After the endorsement meeting has ended, the E&M team will share the meeting recording with all committee members and request they submit their endorsement vote via an off-line voting tool. But the committee endorsement vote (the

Recommendations Group plus the Advisory Group). Committee members will have two (2) business days to submit their votes.

We promote high attendance among voting committee members by engaging them early and often, including providing meeting notices well in advance of scheduled meetings and sending detailed meeting agendas and measure submission information with sufficient time for review. Battelle's enhancements to the E&M process mitigate some of the prior CBE's quorum challenges due to diminished committee engagement and participation. In addition, we acknowledge committee member priorities may change from time-to-time, impacting their ability to meaningfully participate in the E&M process. We employ a policy allowing committee members to be inactive for a given cycle, based on competing priorities, or end a committee member's term early if there is a consistent lack of participation and/or engagement (*see E&M Committee Composition, Roles, and Responsibilities section for more details*).

Endorsement Decision Posted

E&M committee endorsement decisions and associated rationales are posted to the PQM website for three (3) weeks, which represents an Appeals period, during which any interested party may request an appeal regarding any E&M committee endorsement decision.

Appeals

When an appeal is received, we conduct a preliminary review of the appeal to determine its eligibility based on the criteria for the respective endorsement decision. If an appeal is not eligible, we notify the appellant, noting the eligibility criteria not met. However, if an appeal is eligible, an ad-hoc Appeals Committee is convened to review and discuss the appeal, followed by a vote to uphold (i.e., overturn a committee endorsement decision) or deny (i.e., maintain the endorsement decision) the appeal. Consensus is determined to be 75% or greater agreement among members.

Appeal Eligibility Criteria

If a measure's endorsement is being appealed, including an "Endorsed with Conditions" decision, the appeal must:

- Cite evidence of the appellant's interests are directly and materially affected by the measure, and the CBE's endorsement of the measure has had, or will have, an adverse effect on those interests; and
- Cite the existence of a CBE procedural error or information that was available by the cycle's Intent to Submit deadline but was not considered by the E&M committee at the time of the endorsement decision, which is reasonably likely to affect the outcome of the original endorsement decision.

In the case of a measure not being endorsed (formerly a "reconsideration request"), the appeal must be based on one of two rationales:

- The CBE's measure evaluation criteria were not applied appropriately. For this rationale, the appellant must specify the evaluation criteria they believe was misapplied.
- The CBE's E&M process was not followed. The appellant must specify the process step, how it was not followed properly, and how this resulted in the measure not being endorsed.

The Appeals Committee consists of E&M team staff (i.e., technical leaders from Battelle and our partners, IHI and Rainmakers) and all co-chairs from each of the five E&M project committees from the respective endorsement cycle. If additional perspectives are needed, we send ad-hoc requests to the PQM membership. This structure ensures these meetings can be convened quickly and as needed, and the inclusion of E&M staff and committee chairs reduces the risk of duplicative or contradictory discussions. If needed, SMEs may be recruited, as non-voting participants, to support the Appeals Committee discussions. We employ the same SME-recruitment approach as with E&M committees. However, vetting of SMEs will occur amongst the co-chairs of the committees that did not receive an appeal.

To promote transparency and accountability, Appeals Committee meetings are open to the public, and a meeting summary is shared publicly via the PQM website. All Appeals Committee decisions are final.

Final Technical Report

The E&M team develops and publishes a technical report for each project, upon completion. Each technical report includes the following information:

- A summary of the scope of review conducted under the E&M project
- A list of the performance measures submitted and evaluated under the E&M project
- A list of the performance measures endorsed and not recommended for endorsement under the E&M project
- A list of measure concepts submitted during Intent to Submit for measures under the E&M project
- A summary of the public comments received during the E&M process for the E&M project
- A summary of any potential high-priority gap areas for measure developers to consider for future development, identified during the E&M project
- A summary of any major concerns or methodological issues raised during performance measure evaluation of the E&M project.

Harmonization

The current quality landscape contains a proliferation of measures, including some that could be considered duplicative or overlapping and others that measure similar but nonidentical concepts

and/or define patient populations differently. Such duplicative measures and/or those with similar but not identical specifications may increase data collection burden and create confusion or inaccuracy in interpreting performance results for those who implement and use performance measures. Resolving issues around harmonizing measures and handling competing measures is one of the key challenges. Developers/stewards must respond to the questions about harmonization in their measure submission.

ICD-10

The Department of Health and Human Services implemented conversion to ICD-10 coding on October 1, 2015. Further details explaining the changes can be accessed at https://www.hhs.gov/guidance/document/icd-10-general-equivalence-mapping. Battelle requires ICD-10 codes to replace any ICD-9 codes for all new submissions, measures undergoing endorsement maintenance, and measures due for annual update.

Technical Assistance

The E&M project staff provides technical assistance to measure developers and stewards at any time before or during the measure submission process. Contact PQMsupport@battelle.org with any questions about PQM's Measure Evaluation Rubric, how to answer the questions in the form, any technical issues with the online submission process, or anything else.

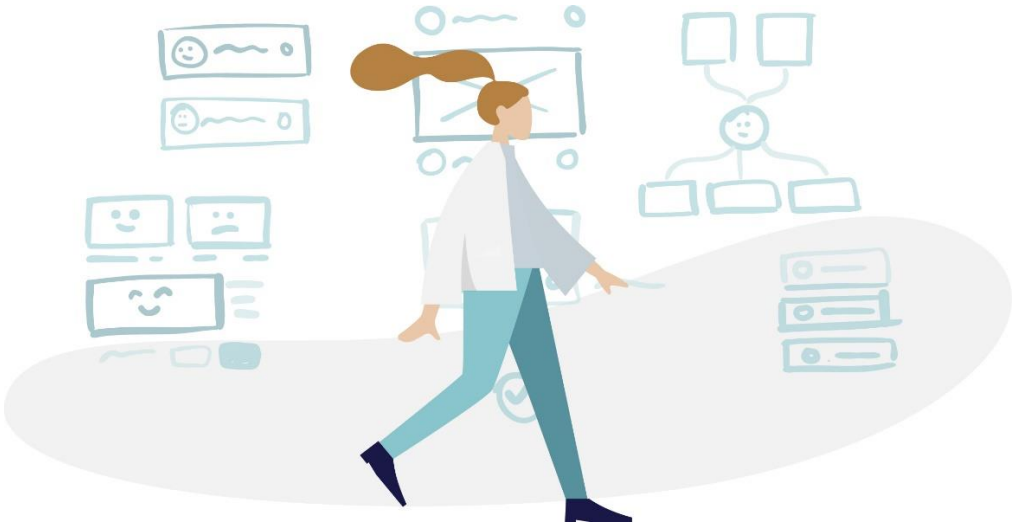
Additional Developer Resources

Battelle engages measure developers extensively to foster discussion and engagement with those who submit measures for E&M review. Materials, webinars, and discussions with measure developers and stewards are intended to promote transparency and a collaborative environment benefiting all interested parties. As a CBE, Battelle cannot engage in measure development. However, each year we host a virtual Measure Developer Workshop, with the intent of engaging measure developers in cutting-edge topics relevant to measurement and E&M. For example, we share recommendations about measure evaluation criteria or testing requirements with measure developers to (1) obtain feedback on the recommendations; and (2) make developers aware of potential changes to future cycles. This more deeply engages measure developers in the refinement of processes and requirements; gives interested parties a "heads-up" as to what is coming at every stage, with the intention of improving overall openness to the changes; and contributes to consensus-building by providing an opportunity for us to gather in-depth input and recommendations on improvements to the process.

In addition, Battelle staff, who have measurement expertise, can assist developers through the submission and review process. The Battelle team can serve as a resource to developers through deep and nuanced understanding of the quality measure lifecycle, the tools and resources required to develop measures, the underlying measure science that guides measure development, and the time and resource constraints that impact measure development.

Maintenance of Endorsement

Maintenance of endorsement encompasses several processes: (1) annual updates to measure specifications of endorsed measures, (2) evaluations for endorsement maintenance (3) emergency/off-cycle reviews, (4) analysis and guidance for methodological and technical challenges, and (5) education and technical assistance to measure developers on endorsement maintenance activities. Once a measure is endorsed, it will enter a 5-year maintenance cycle, at which time the measure is resubmitted to Battelle for PQM endorsement review. However, prior to the 5-year maintenance review, at three (3) years since the measure's endorsement, developers/stewards provide an annual update indicating whether any changes to the measure specifications are needed. The developer/steward may also submit an attestation if no changes are required. Once the three-year annual update is submitted, we will review to confirm if any changes indicated require the measure to be submitted for endorsement review before the five-year maintenance cycle.



Annual Updates

Every year, when an endorsed measure is not being re-evaluated for continued endorsement, measure stewards have the option to submit a status report of the measure specifications to Battelle. This report either reaffirms the measure specifications remain the same as those at the time of endorsement or last update or outline any changes or updates made to the endorsed measure, including the purpose for the changes.

If changes occur to a measure at any time in between the measure's last endorsement review and its scheduled maintenance endorsement review, the measure steward is responsible for informing Battelle immediately by submitting a status report. An early maintenance review is conducted if the changes materially affect the measure's original concept or logic (see *Emergency Review/Off-Cycle Reviews* below).

Emergency Review/Off-Cycle Reviews

An emergency or off-cycle review is a formal measure evaluation and endorsement consideration occurring prior to the previously scheduled maintenance of endorsement date. An early maintenance review follows the same process as a maintenance of endorsement evaluation.

An emergency or off-cycle review is triggered by a variety of ways:

- Request by a developer/steward due to a material change to an endorsed measure during an annual update. A material change is defined as any significant modification to the measure specifications, which significantly affect the measure results such as:
    - Changes to the population being measured (e.g., changes in age inclusions, changes in diagnoses or other inclusion criteria, changes in excluded populations, from one type of insured population to all-payer population);
    - Changes to what is being measured (e.g., changes in target values like blood pressure or lipid values);
    - Inclusion of new data source(s); or
    - Expansion of the level or changing unit of analysis or care setting(s) (e.g., adding clinician-level to a measure currently endorsed at practice-level).
- Request by an interested party because of a perceived unintended negative consequence associated with the measure, a change in the clinical guideline driving the measure, or a significant implementation issue. The interested party may be a measure developer/steward, E&M committee member, or any other type of interested party.

Battelle restricts the scope of the emergency or off-cycle review to the immediate issue (i.e., concern with the measure's evidence) and not an all-encompassing review. An early maintenance review can be requested by any party, if there is adequate, high quality, and consistent evidence to justify the review. To initiate the review, the interested party must send an email to PQMSupport@battelle.org with the subject "Emergency/Off-cycle Review Requested," which alerts the E&M project team. The project team and respective E&M committee co-chairs review the request to see whether it is significant and emergent; for example, if the clinical practice underlying the measure is causing harm to patients directly or as a result of an unintended consequence. If deemed significant and emergent, the project team notifies the developer/steward (if they are not the requester of the emergency review) and pulls the measure off its maintenance cycle to be reviewed by the E&M committee during the next immediate cycle.

The E&M team recruits additional SMEs, as needed, ensuring an appropriate combination of perspectives, from PQM and from our partners, IHI and Rainmakers. The E&M committee determines whether the measure needs immediate attention, such as a change to the specifications, and shares this information with the measure developer/steward. If the change is not feasible, the committee may decide to remove the measure's endorsement. If the measure does not need immediate attention, the measure developer/steward should document the issue

for consideration in the next round of full review. The E&M team informs the requester of the final decision with justification.

Appendix A: Quality Measure Developer and Steward Agreement

Each candidate measure or set of measures has a measure steward who assumes responsibility for the submission of the measure to Battelle for potential endorsement. The measure steward is responsible for making necessary updates to the measure and informing Battelle about any changes made to the measure. In addition, the measure steward is responsible for providing the required measure information during the measure maintenance process:

- The measure steward organization is required to identify a single point of contact who will be notified of any upcoming maintenance deadlines or requirements related to the endorsed measure(s). If there is a change in point of contact, then the steward should notify Battelle of the new point of contact.
- Stewards may be contacted by PQM members or other members of the public with inquiries about specifications, updates, and implementation of the endorsed measure(s)
- Stewards are also responsible for maintaining measure details and specifications on any publicly available website.

Each steward who submits a fully specified and tested measure to Battelle must submit a completed and signed Quality Measure Developer and Steward Agreement (QMDSA) on or before the project's measure submission deadline in order for the measure to be considered by the Committee. The agreement is between Battelle and the measure steward and only shared between these parties:

- For new measure stewards, the QMDSA should be accompanied by the completed addendum, in which the steward must list all the measures (measure number and measure title) being submitted for review.
- For existing measure stewards, only a signed addendum is needed and will be appended to the existing MSA; a new MSA is not required. Contact Battelle project staff to obtain the addendum.

Only one QMDSA is necessary per measure steward. If the steward is a governmental organization, a QMDSA is not required. For more information about how to complete the QMDSA, please see the QMDSA Submission Instructions.

Battelle will work with all measure stewards to transition to this QMDSA. Those who have measures up for maintenance or wish to add additional measures to their current QMDSA will need to complete an Additional and Maintenance Measures Form. Each QMDSA will stand for five (5) years from its effective date.

The QMDSA and Additional and Maintenance Measures Forms are contractual agreements that must be signed by Battelle and any measure steward that is submitting one or more measures to be evaluated for endorsement.

Appendix B: Personal/Organizational Disclosure of Interest Form

1. Name:

   Organization Affiliation:

   Committee Name:

   Year:

NOTE: *This form will be renewed annually, please denote the year this disclosure will cover (ex. 2023) If you fill this form out in August of any given year, you will need to submit a new form January of the following year.*

2. Describe any personal or organizational relationships subject to disclosure. If None, check here: ☐

3. Describe any personal or organizational financial interests subject to disclosure. If None, check here: ☐

4. Electronic Certification
   By executing this Electronic Certification, I certify that I have reviewed the Personal/Organizational Disclosure of Interest Form, and the information given above is true to the best of my knowledge.

Name: _____ **Signature:** _____

Date: _____

All persons and organizations must be free of any conflicts of interest financially for this effort. If at any time you believe that a potential or actual conflict exists, you must notify Battelle immediately. “Conflict of Interest” means because of other activities or relationships with other persons or organizations you are unable or potentially unable to (1) render impartial assistance or advice; (2) perform due to the impairment of or the possibility of the impairment of your objectivity; or (3) perform because you have or might acquire an unfair competitive advantage.

Appendix C: Measure Disclosure of Interest Form

1. Name:

   Organization Affiliation:

   Committee Name:

   Cycle (ex. Fall 2025):

2. Describe any personal or organizational measure conflicts. If None, check here: ☐
   a. Measures under review:

CBE #	Measure Title	Measure Developer/Steward
####	[insert title]	[insert developer and steward]
####	[insert title]	[insert developer and steward]
####	[insert title]	[insert developer and steward]

      i. If you have worked as an employee, collaborator, or consultant of the measure developers/stewards listed *OR* contributed to the development of the measures listed, in any capacity, in the past five (5) years, check here: ☐

   b. Competing Measures:

CBE #	Measure Title	Measure Developer/Steward
####	[insert title]	[insert developer and steward]
####	[insert title]	[insert developer and steward]
####	[insert title]	[insert developer and steward]

      i. If you have worked as an employee, collaborator or consultant of the measure developers/stewards listed *OR* contributed to the development of the measures listed, in any capacity, in the past five (5) years, check here: ☐

   c. If you checked either box under 2a. or 2b. above, please provide a detailed description of the involvement. (Include CBE ID number, Measure Title, Cycle, and Steward Name:)

(continued on next page)

3. Electronic Certification

   By executing this Electronic Certification, I certify that I have reviewed the Measure Disclosure of Interest Form, and the information given above is true to the best of my knowledge.

Name: _____ **Signature:** _____

Date: _____

All persons and organizations must be free of any conflicts of interest for measures under this effort. If at any time you believe that a potential or actual conflict exists, you must notify Battelle immediately. "Conflict of Interest" means because of other activities or relationships with other persons or organizations you are unable or potentially unable to (1) render impartial assistance or advice; (2) perform due to the impairment of or the possibility of the impairment of your objectivity; or (3) perform because you have or might acquire an unfair competitive advantage.

Appendix D: PQM Measure Evaluation Rubric

Note: Rubric items correspond to items in the measure submission form and provide the information needed to evaluate each of the five Rubric domains.

The requirements for initial and maintenance measure endorsement are indicated as, "*[For initial endorsement]*" or "*[For maintenance],*" within each domain of PQM Measure Evaluation Rubric. If neither distinctions are listed for a rubric requirement, then it applies to both initial and maintenance endorsement.

The PQM Measure Evaluation Rubric does not include must-pass criteria, nor algorithms for assigning a rating. Rather, the PQM Measure Evaluation Rubric guides reviewers to a rating of "Met", "Not Met, but Addressable", or "Not Met" based on the criteria listed for each. As part of its continuous quality improvement of the E&M process, Battelle considers whether changes to the domains, criteria, and/or additional guidance, such as an algorithm, are needed.

Importance
Attach a logic model depicting the relationship between structures and processes and the desired outcome.
Summarize evidence of measure importance from the literature linking the structure/process/intermediate outcome to the outcome
[For initial endorsement] If implemented, what is the measure's anticipated impact on important outcomes?
[For maintenance] Provide evidence of performance gap or measurement gap by providing performance scores on the measure as specified (current and over time) at the specified level of analysis
Explain why existing measures/quality improvement programs are insufficient for addressing this health care need?
Provide evidence the target population (e.g., patients) values the measured outcome, process, or structure, and finds it meaningful. Describe how and from whom you obtained input.

Not Met:

- Evidence is about something other than what is measured OR
- Empirical evidence submitted without literature review or grading OR
- Empirical evidence includes only selected studies from the literature review[2] OR

[2] A literature review could include a systematic review, clinical practice guidelines, observational studies, case studies, etc. The purpose of the literature review is to identify relevant studies to support the measure's logic model. Developer/stewards should provide a summary of the evidence for the committee's consideration. An evaluation of the quality of evidence should also be conducted. Often clinical practices guidelines conduct systematic reviews. If a literature review is not possible, a rationale as to why would be considered by the committee.

- Evidence is not graded high quality or strong recommendation OR
- Literature review conclusion is that consistency is low or controversial; moderate/high certainty that the net benefit (i.e., improved outcomes, adverse events and/or costs avoided due to the measure's anticipated impact) is null or small; or grade of weak OR
- There is low confidence/certainty that there is an adequate business case[3] (the anticipated impacts of the measure on patient outcomes and/or costs/resource use justify the measure and its use), where "adequate"=there is a net benefit to measurement OR
- There is low confidence/certainty that there is evidence of a performance gap, as determined by variation in performance or less than optimal performance for the overall target population and/or subpopulations OR
- There is no description of other existing measures or programs or no search conducted to identify other existing measures or programs OR
- Proposed measure has the same measure focus and target population as existing measures and offers no advantage in terms of addressing disparities, feasibility, potential use, or scientific acceptability OR
- Patient input does not support the conclusion that the measured outcome, process, or structure is meaningful or it does so with a low degree of certainty.

Not Met but Addressable:

- Criterion is not met (see above), but the reviewer can identify changes to specifications that may strengthen the measure's importance such that the criterion could be met.

Met:

- Literature review concludes with at least moderate certainty that a net benefit (i.e., improved outcomes, adverse events and/or costs avoided due to the measure's anticipated impact) is at least moderate AND
- There is at least moderate confidence/certainty that there is an adequate business case (i.e., the anticipated impacts of the measure on patient outcomes and/or costs/resource use justify the measure and its use), where "adequate"=there is a net benefit to measurement AND
- There is at least moderate confidence/certainty that there is evidence of a performance gap, as determined by variation in performance or less than optimal performance for the overall target population and/or subpopulations AND

[3] For more information on how to consider the business case for a measure, please refer to the CMS Measure Management System Blueprint

- Description of existing measures or programs justifies the proposed measure's focus among the proposed measure's target population and/or the proposed measure is superior[4] to identified related or competing measures AND
- Description of patient input supports the conclusion that the measured outcome, process, or structure is meaningful with at least moderate certainty.

Feasibility
[For Initial Endorsement] Describe the feasibility assessment showing you considered the people, tools, tasks, and technologies necessary to implement this measure. If an eCQM, please attach your completed eCQM Feasibility Scorecard.
Describe how the feasibility assessment informed the final measure, indicating any decisions made to adjust the measure in response to data availability.
Describe any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm.

Not Met:

- Feasibility assessment not systematically conducted or described OR
- Long-term or no path is specified to support routine and electronic data capture with an implementable data collection strategy.

Not Met but Addressable:

- Criterion is not met (see above), but the reviewer can identify changes to specifications that may improve feasibility such that the criterion could be met.

Met:

- Near-term paths are specified to support routine and electronic data capture with an implementable data collection strategy OR
- Required data are routinely generated and used during care, required data are available in EHRs or other electronic sources, and the data collection strategy can be implemented.

[4] Measure developers/stewards must document why the proposed measure is superior to any identified and/or competing measures and should include any literature used to support this position. For instance, clinical practice guidelines supporting the proposed measure do not support any existing measures identified; or the proposed measure's intentions vary across programs/payors, which requires the measure to be distinct from other existing measures; or the proposed measure captures a target population at higher risk such at the use of the proposed measure may close care gaps for a higher-risk population.

Scientific Acceptability
Describe the data or sample used for testing (include dates, source). If you used multiple data sources for different aspects of testing (e.g., reliability, validity, risk adjustment), identify how the data or sample are different for each aspect of testing.
Provide descriptive characteristics of measured entities included in the analysis (e.g., size, location, type). If you used a sample, describe how you selected entities for inclusion in the sample.
Identify the number and descriptive characteristics (e.g., age, sex, race, diagnosis), of the unit of analysis, for example, patient, encounter or episode, separated by level of analysis and data source. If you used a sample, describe how you selected the patients for inclusion in the sample. If there is a minimum case count used for testing, you must reflect that minimum in the specifications.
If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), please identify how the data or sample are different for each aspect of testing.
Select the level of reliability testing conducted. ☐ Patient or Encounter-Level (e.g., inter-abstractor reliability) ☐ Accountable Entity Level (e.g., signal-to-noise analysis)
For each level of reliability testing conducted, describe the method of reliability testing and what it tests.
Provide the statistical results from each level of reliability testing conducted and at the measure's level of analysis (e.g., clinician, health plan, facility).
Provide your interpretation of the results in terms of demonstrating reliability (i.e., How do the results support an inference of reliability for the measure?)
Select the level of validity testing conducted. ☐ Patient or Encounter-Level (e.g., sensitivity and specificity) ☐ Accountable Entity Level (e.g., criterion validity)
Select the type of validity testing conducted. ☐ Empirical validity testing (e.g., data element testing, empirical testing of measure score) ☐ Systematic assessment of face validity of performance measure score as an indicator of quality or resource use (i.e., the score is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance).
For each level of testing conducted, describe the method of validity testing and what it tests.
Provide your interpretation of the results in terms of demonstrating validity (i.e., How do the results support an inference of validity for the measure?)
Check all methods used to address risk factors. ☐ Statistical risk model with risk factors (___Specify number of risk factors) ☐ Stratification by risk category (___Specify number of categories) ☐ Other (___Specify) ☐ No risk adjustment or stratification

Scientific Acceptability
Attach a conceptual model illustrating the pathway between patient risk factors (social, functional status-related, and clinical factors), quality of care, and the measured outcome. Explain the rationale for the model.
Provide descriptive statistics on the distribution across the measured entities of the risk variables identified in the conceptual model.
If using statistical risk models, provide detailed risk model specifications (query or algorithm), including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.
Detail the statistical results of the analysis used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.
Provide the approach and results of calibration and discrimination testing. Describe any over- or under-prediction of the model for important subgroups.
If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate there is no need to control for differences in patient characteristics (i.e., case mix) to achieve fair comparisons across measured entities.

Not Met:

Sampling

- Sampling is used and sampling strategy is not determined by the measure's analytic unit OR sample does not represent variety of entities whose performance will be measured OR sample does not include adequate numbers of units of measurement for the selected statistical method OR

For Patient or Encounter Level Reliability[5]

- Internal consistency < 0.7 OR
- Inter-rater agreement < 0.4 OR
- Test-retest reliability (Intraclass correlation or Pearson correlation) < 0.5 OR
- Linear relationship < 0.6 OR

For Accountable Entity Level Reliability[5,6]

- Signal to noise/Inter-unit Reliability < 0.6 OR

[5] Reliability thresholds were established by the Scientific Methods Panel and confirmed at the June 14, 2022 advisory meeting.

[6] For accountable entity level reliability testing, the associated thresholds apply to the accountable entity (e.g., facility, clinician, health plan), not the mean or median across all entities.

- Split-half reliability (ICC) < 0.6 OR

Validity

- Face validity is inadequate[7] OR is the only type of validity discussed and the measure is undergoing maintenance review OR
- Reviewer determines the methodology to assess validity is inadequate/inappropriate[8] OR the analytic approach is inadequate/inappropriate OR
- Reviewer disagrees with the assertion that the measure can distinguish quality with limited or no threats to validity present OR

Risk Adjustment

- Factors in the risk model do not influence the measured outcome OR are not present at the start of care OR the risk model includes factors that are associated with differences or inequities in care without sufficient rationale based on the conceptual model OR
- Analysis does not demonstrate:
    - Variation in prevalence of risk factors across measure entities AND
    - Contribution to unique variation in the outcome AND
    - Impact of risk adjustment for providers at high or low extremes of risk OR
    - Results do not demonstrate acceptable model performance.

Not Met but Addressable:

- Criterion is not met but the reviewer can identify:
    - Improvements to the sampling methodology OR
    - Changes to the methodology/analytic approach that could improve assessment of reliability OR
    - Changes to the methodology/analytic approach that could improve assessment of validity OR
    - Changes to the specifications that could improve validity and/or address threats to validity OR

[7] Face validity is accomplished through a systematic and transparent process, in which developers/stewards disclose identified relevant experts (e.g., clinicians, accountable entity representatives, those [patient, caregivers] with lived experience) and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

[8] As part of the validity testing methodology, developers/stewards should empirically assess, as appropriate, the impact of missing data and/or measure exclusions.

- Changes to the risk model that could improve model appropriateness or performance.

Met:

Sampling

- If a sample is used, the sampling strategy is determined by the measure's analytic unit AND sample represents the variety of entities whose performance will be measured AND sample includes adequate numbers of units of measurement for the selected statistical method AND

For Patient or Encounter Level Reliability[5]

- Internal consistency $\geq$ 0.7 OR
- Inter-rater agreement $\geq$ 0.4 OR
- Test-retest reliability (ICC or Pearson correlation) $\geq$ 0.5 OR
- Linear relationship $\geq$ 0.6 AND

For Accountable Entity Level Reliability[5,6]

- Signal to noise/Inter-unit Reliability $\geq$0.6 OR
- Split-half reliability (ICC) $\geq$ 0.6 AND

Validity

- Face validity is adequate[7] and the measure is undergoing initial review OR
- Reviewer determines methodology employed[8] is adequate and the analytic approach presented is appropriate and thorough AND
- Reviewer determines results of empirical testing adequately demonstrate that the measure is valid AND
- Reviewer determines the interpretation of the empirical results supports an inference of validity AND

Risk Adjustment

- Factors in the risk model influence the measured outcome AND are present at the start of care AND the risk model does not include factors that are associated with differences or inequities in care unless justification provided based on the conceptual model AND
- Analysis demonstrates:
    - Variation in prevalence of risk factors across measured entities AND
    - Contribution to unique variation in the outcome

- Impact of risk adjustment for providers at high or low extremes of risk AND
- Results demonstrate acceptable model performance.

Equity*
Describe how this measure contributes to efforts to address inequities in health care. Provide a description of your methodology and approach to empirical testing of differences in performance scores across multiple sociocontextual variables (e.g., race, ethnicity, urbanicity/rurality, SES, gender, gender identity, sexual orientation, age). Provide an interpretation of the results, including interpretation of any identified differences and consideration of negative impact or unintended consequences on subgroups.

**The Equity domain is optional, as Battelle recognizes some measures are not designed to advance health equity. Battelle continues to explore this, but to align with national priorities, Battelle encourages developers and stewards to address this domain, if and when possible.*

Not Met:

- Reviewer determines equity is not sufficiently assessed OR the measure does not contribute to efforts to address inequities in health care.

Not Met but Addressable:

- Criterion is not met but reviewer can identify changes to the assessment of equity OR changes to the measure specifications that would address inequities in health care.

Met:

- Reviewer determines sufficient assessment of equity was conducted (i.e., methodology provided, differences in scores tested across multiple categories, and interpretation of results) AND the measure contributes to efforts to address inequities in health care.

Use and Usability
[For initial endorsement] Check all planned uses and provide the name of the program and sponsor, URL, purpose, geographic area and percentage of accountable entities and patients included, and level of analysis and care setting. ☐ Social Security Act modifications under the Patient Protection and Affordable Care Act and related accountability applications ☐ Quality Payment Program (QPP) Merit-based Incentive Payment System (MIPS) and Qualified Clinical Data Registries (QCDRs)

Use and Usability
☐ Specialty society clinical data registrations ☐ Certification programs ☐ Employer insurance plans ☐ Medicaid ☐ Other use:
[For maintenance review] Check all current uses. ☐ Social Security Act modifications under the Patient Protection and Affordable Care Act and related accountability applications ☐ QPP MIPS and QCDRs ☐ Specialty society clinical data registrations ☐ Certification programs ☐ Employer insurance plans ☐ Medicaid ☐ Other (specify):
What are the actions measured entities can take to improve performance on this measure? How difficult are those actions to achieve?
[For maintenance only] Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how you obtained feedback.
[For maintenance only] Describe how you considered the feedback when developing or revising the measure specifications or implementation, including whether you modified the measure and why or why not.
[For maintenance only] Discuss any progress on improvement (trends in performance results, including performance among sub-populations, if available, number and percentage of people receiving high-quality health care, geographic area, number and percentage of accountable entities and patients included). If use of the measure demonstrated no improvement, provide an explanation.

Not Met:

For initial endorsement

- There is no plan for use in at least one accountability application after initial endorsement but before the measure's first maintenance review OR
- Performance scores do not yield actionable information that can be used to improve performance among measured entities.

For maintenance

- The measure is not currently in use in at least one accountability application OR
- Performance scores do not yield actionable information that can be used to improve performance among measured entities OR
- Reviewer determines, based on the information provided regarding feedback on measure performance, the measure is not usable.

Not Met but Addressable:

For initial endorsement and maintenance

- Criterion is not met (see above), but the reviewer can identify changes to specifications that may strengthen the measure's ability to yield actionable information or usability.

Met:

For initial endorsement

- There is a plan for use in at least one accountability application after initial endorsement but before the measure's first maintenance review AND
- Performance scores yield actionable information that can be used to improve performance among measured entities.

For maintenance

- The measure is currently in use in at least one accountability application AND
- Performance scores yield actionable information that can be used to improve performance among measured entities.
- Reviewer determines, based on the information provided regarding feedback on measure performance, the measure is usable.

Appendix E: Guidance to Make Submissions 508 Compliant (required)

Battelle ensures all public facing materials are 508 compliant. Section 508 of the Rehabilitation Act of 1973 (29 U.S.C. 794d), as amended by the Workforce Investment Act of 1998 and the Architectural and Transportation Barriers Compliance Board Electronic and Information (EIT) Accessibility Standards (36 CFR part 1194), ensures those with disabilities have equal access to government information as contained on information and communications technology (ICT), and thereby to the government employment, programs and services to which all citizens are entitled. The following steps should be taken during the measure submission process to maintain Section 508-compliance:

- Creating tables with row and column headers and proper reading order. Tables must be properly created to be Section 508-compliant. The table feature of the software must be employed, rather than tabs and drawn lines. Row and column headers must be identified as such. Tables must be set up such that the reading order is left to right and top-down in order to be read correctly by read-aloud software. Tables should avoid any merging or splitting of cells. Table rows should not split/break across pages. Repeat the column and row headers to avoid merging cells and issues with splitting/breaking rows across pages.
- Providing alternative text (alt-text), a text equivalent describing images, graphics, and exhibits that can be used by text-to-speech programs. Developers/stewards should provide alt-text of the image, chart, or graphic, describing all the information important for understanding the image for the visually impaired web user. All of the relevant information in charts, graphs, and diagrams should be included in the alt-text. Images or graphics that are added for design or layout only (i.e., add no meaning to the document) can be described briefly, as in “bullet” or “empty cell.”
- Using color appropriately. There must be enough color contrast in graphics to prevent those individuals with color vision deficiencies from having problems understanding the graphic. Color alone cannot be used to convey information or meaning.
- Creating hyperlinks using a description of the link destination rather than vague or confusing text such as “click here.”

The E&M team provides a checklist of 508 compliance criteria for developers/stewards to consider when submitting measures to Battelle (see below).

508 GUIDANCE CHECKLIST

Version: 1.0; Generated: 14 April 2023

These guidelines apply to all parts of your measure submission including all fields and attachments used within the measure submission forms.

Text

- ☐ Is all my text black, not using any other colors?
- ☐ Am I reserving underlined text for **hyperlinks** only and creating emphasis using *italic,* **bold,** and ***bold-italic*** text instead of using underlining?
- ☐ Am I avoiding multiple hard and soft returns?
- ☐ Are all my hyperlinks working, linked to their correct destination, and using a distinct style to set them off from regular text?
- ☐ Do all my bulleted or numbered lists use the built-in bulleting or numbering options?

Tables

- ☐ Are my tables actual tables and **not** images or screenshots of a table?
- ☐ Am I using a table creation tool or a attaching a Word *Table Design Style* table?
- ☐ Am I repeating the column and row headers in individual cells to avoid merged table cells?
- ☐ Do my empty table cells contain a symbol like * with the note: **Cells intentionally left empty,* at the bottom outside of my table?
- ☐ Did I write a brief description of what the table conveys using the Table Caption option?
- ☐ Is the table converted to paragraph text if it is too long to fit all of one column on a single page and flows over to the next page?
- ☐ Does my attached Word table have *Allow row to break across pages* turned **off** for all rows and *Repeat as header row at the top of each page* turned **on** for the first row?

Images, Figures, Graphs, Charts, and Pictures

- ☐ Do my images include clear concise alt-text descriptions of what they represent using the image caption option, or Edit/Alt-text option for Word attachments?

Appendix F: Measure of Consensus

- Variance is used as a metric to assess disagreement (lack of consensus)
- However, variance alone is insufficient when comparing different sizes of groups or groups with different means
- The measure of consensus is the complement of the index of disagreement, which is based on the variance of the responses scaled by the total available range of variance conditional on the mean response.

Number of respondents	Endorse	Endorse with Conditions	Not Endorse/Remove Endorsement	Measure of Consensus
40	0.000	0.250	0.750	1.00000
40	0.125	0.125	0.750	0.99429
20	0.000	0.250	0.750	1.00000
20	0.125	0.125	0.750	0.95170
40	0.125	0.750	0.125	0.99707
20	0.150	0.750	0.100	0.97065
40	0.250	0.000	0.750	0.94527
20	0.250	0.000	0.750	0.95110
40	0.500	0.000	0.500	0.81789
20	0.500	0.000	0.500	0.80713

Measure of Consensus (far-right column) = 1 minus the Index of Disagreement

Index of Disagreement (not shown) = Response variance / Total available range of variance

At 0.95000 for the Measure of Consensus (far-right column), at least 75% of respondents are in agreement (i.e., 75% of the respondents [blue shading] voted to Endorse, Endorse with Conditions, or Not Endorse/Remove Endorsement). As the response variance increases, the more disagreement there is amongst respondents, and the Measure of Consensus (far-right column) decreases. Perfect agreement would mean there is zero variance, which may be insurmountable given the differences of opinions, expertise, and/or experience of respondents. Therefore, the E&M process employs a Measure of Consensus of 95%, which corresponds to a consensus threshold of 75% amongst respondent votes.

References

M. A. Rahem and M. Darrah, “A geometric approach for computing a measure of consensus for groups,” International Mathematical Forum, vol. 11, pp. 961–973, 2016.

M. A. Rahem and M. Darrah, "Using a Computational Approach for Generalizing a Consensus Measure to Likert Scales of Any Size." International Journal of Mathematics and Mathematical Sciences, 2018.

Y. Akiyama, J. Nolan, M. Darrah, M. A. Rahem, and L. Wang, “A method for measuring consensus within groups: an index of disagreement via conditional probability,” Information Sciences, vol. 345, pp. 116–128, 2016.

Y. Tsuchiya and N. Hiramoto, “Measuring consensus and dissensus: a generalized index of disagreement using conditional probability,” Information Sciences, vol. 439/440, pp. 50–60, 2018.

Appendix G: Public Comments and Battelle Responses

Overview

The E&M Guidebook was posted on the E&M webpage of the PQM website for public comment on July 1, 2023 for 30 calendar days. During this commenting period, Battelle received 44 total comments from 25 organizations and 10 individuals. Prompts provided for public response included:

- What general comments do interested parties have on the report?
- What are comments do interested parties have about the various sections of the E&M Guidebook, including but not limited to:
    - E&M Committee Structure
    - Conflict of Interest Policy
    - Submitting Measures to Battelle
    - E&M Processes
    - Quorum and Voting
    - Appeals
    - PQM Measure Evaluation Rubric

All comments received have been posted to the PQM website.

Comment Themes

Below is a summary of key themes emphasized by commenters and Battelle's response.

General Comments

Fifteen comments received did not pertain to a specific section of the E&M Guidebook. Rather, commenters expressed the need for rewording, clarity, more detail, formatting, definitions, etc., overall. Battelle thanks commenters for this feedback and has provided updates to various sections to provide more clarity and consistency throughout the document.

In addition, Battelle received supportive comments with respect to the overall E&M process, including support for/of the:

- Retirement of the CSAC committee
- Increase in the amount of independent committee measure reviews
- Revised role of the SMP
- Increase in the consensus threshold to 75%
- Recruitment of SMEs into the E&M process
- Use of the Novel Hybrid Delphi and Nominal Group (NHDNG) technique
- Addition of the Equity criterion within the PQM Measure Evaluation Rubric

Battelle thanks the commenters for their support.

E&M Committee Structure and Conflict of Interest

Thirty comments pertained to the E&M project committee structure. Most of these comments requested more transparency and detail with respect to identifying SMEs and their role within the E&M process, including if the COI policy applies to SMEs. Several comments raised concern that the reduced number of committees may increase the burden on committee members and the conditions listed under the End-of-Life committee should not be, as patients with those conditions (e.g., end-stage renal disease) "…live 20+ years." Lastly, the E&M team also considered comments received for the Pre-Rulemaking Recommendations and Measure Set Review (PRMR-MSR) Guidebook, which held a concurrent public comment period. Several PRMR-MSR comments requested more transparency in

committee selection, namely how randomization would occur when seating committee members to the Advisory vs. Recommendation Group.

Battelle appreciates the request from commenters for more detail and transparency regarding the SME selection process and the role of SMEs within the E&M process. Battelle has provided more detail, noting SMEs are non-voting members and are invited to committee endorsement meetings to provide context and relevance to the measure submission, including answering questions posed by the committee. For the Fall 2023 cycle, Battelle will establish whether SMEs/additional expertise is needed after Intent to Submit (ITS). If additional expertise is needed, Battelle will first identify if expertise resides within one of the other E&M committees. If expertise is still lacking, Battelle and its partner, IHI, will recruit SMEs from our combined networks, and selection would be made with input from the respective committee co-chairs. Any SME recruited to participate will be held to the same conflict of interest policy as committee members. For future cycles, Battelle will consider establishing a pool of SMEs across various clinical (e.g., nephrologists, primary care providers) and methodological areas (e.g., psychometricians). These SMEs will also have term limits, which may be the same as committee members, and will also be non-voting participants during endorsement meetings.

With respect to the burden to committee members, Battelle intends to streamline the E&M process by implementing the NHDNG technique, which will reduce iterative committee/panel reviews and focus committee discussions on key issues/areas, especially where consensus is lacking. For the Fall 2023 cycle, Battelle anticipates committees will review no more than eight measures per committee, unless an off-cycle or emergency review is needed. Battelle has also revised the name of the End-of-Life committee to Advanced Illness and Post-Acute Care to better reflect the conditions and measures within this project's portfolio.

Lastly, to align with the PRMR-MSR response to the randomization approach for seating the committee subgroups, for the first year (Fall 2023 and Spring 2024), Battelle will use a randomization generator to apply a 1, 2, or 3-year term to individuals with each roster category. Those members who were allocated to a 1-year term will be included in the Recommendations Group. At the end of their term, those the 2-year term members will be moved to the Recommendations Group and their seats will be open for nominations next year (except for co-chairs). This process will continue so that all new members must serve on the Advisory Group for two years before they can participate in the Recommendations Group during their last year. This ensures each committee member will have an opportunity to serve on the Recommendations Group after learning how the process is run by serving on the Advisory Group at the beginning of their term (except of the first year).

Submitting Measures to Battelle

Fourteen comments were received requesting for more information about when and how to submit measures for endorsement review via the Submission Tool and Repository (STAR). In addition, a few comments noted the searchable database on the PQM website does not currently have full measure submission materials of all measures previously reviewed by the prior CBE. Several comments requested for more guidance on what is required for both initial endorsement and maintenance endorsement review and for more clarity on where measure concepts would be categorized; for example, which committee/project will evaluate patient-reported outcome performance measures (PRO-PMs) and post-acute care measures.

Battelle appreciates these comments and the patience of developers and stewards as Battelle launches the Fall 20223 ITS and Full Measure Submission (FMS) process. Battelle has noted in this Guidebook the STAR will be available for the Fall 2023 cycle and has added links to the Microsoft Word versions of the ITS and FMS forms, which have also been posted to the PQM website. To further detail the measure submission process, Battelle hosted an informational webinar on September 15 about the measure submission process via STAR.

Regarding the full measure specifications of previously submitted measures to NQF, Battelle will be contacting all measure stewards to establish a new steward agreement with Battelle. At which point, Battelle will post the detailed measure specifications for these measures to the PQM website.

With respect to the requirements for initial and maintenance measure endorsement, this distinction resides within the PQM Measure Evaluation Rubric. However, this was not stated in the body of the Guidebook. Therefore, Battelle has noted within the body of the Guidebook, specific expectations for initial vs. maintenance endorsement can be found within the PQM Measure Evaluation Rubric.

Lastly, the measure recategorization from the prior CBE projects to the five new projects is based on the patient journey. If expertise is lacking within the project committee for evaluating certain measures, like PRO-PMS, Battelle will recruit SMEs, using the approach noted above. Battelle has also updated Table 1 with additional areas covered within each project and provided more measure examples.

E&M Processes

Thirty-one comments were received regarding the E&M process. Commenters requested clarification on how public comments will be shared with developers and additional opportunities for public comment with respect to further changes to the Guidebook. Some commenters requested to not have public comment deadlines end on a weekend or holiday and to not have the public comment periods between the E&M and PRMR-MSR processes overlap. One commenter noted the E&M Appeals Period is not a true public comment period, where stakeholders can share their opinions and perspectives without any limitation. Lastly, several comments also requested for more information with respect to the “Endorsed with Conditions” endorsement decision, specifically, what conditions would be included in this category, how will measure implementers, like CMS, interpret/consider this endorsement decision, and what happens if to subsequent endorsement decisions if conditions are not met?

Battelle thanks the commenters and has revised the Guidebook accordingly. Regarding the E&M public comment periods, all comments are posted to the PQM website within two business days of receipt for developers, stewards, and any interested party to view. Devleoper and stewards, therefore, have access to public comments via the PQM website. For the first comment period prior to the endorsement meeting, Battelle encourages developers and stewards to monitor the comment pages of their respective measure(s). Battelle will not be asking developers and stewards to provide responses prior to the endorsement meeting. Rather, developers may address these comments during designated time points at the endorsement meeting (e.g., opening remarks, final responses). Battelle also recognizes the concern with the overlapping public comment periods with PRMR-MSR and will consider potential changes to timelines after the first year. Battelle has noted the public comment periods will end on a business day. Lastly, Battelle appreciates the comment about the Appeals Period not being true public

comment period but respectfully disagrees. E&M public comment periods have a particular intent, which helps to focus the comments being submitted. The intent of the Appeals Period is not to limit perspectives or opinions but rather provide an opportunity for any member of the public to submit a comment in the form of an appeal. Submitting comments focused on the intent of the commenting period also occurs during the first public comment period. During the pre-endorsement comment period, Battelle requests for public comments regarding any concerns with or support for the measure(s) under endorsement review.

With respect to the "Endorsed with Conditions" endorsement decision, Battelle thanks the commenters for their requests for more clarification with this category. Battelle has included examples of the types of recommendations that may be considered for the "Endorsed with Conditions" category (e.g., providing updated testing across a larger population and/or different level of analysis), and noted the committee should also consider what is feasible and appropriate for the developer/steward to accomplish. Battelle also noted several non-negotiable conditions (e.g., inappropriate methodology or testing approach applied to demonstrate reliability or validity). When considering implementation of "Endorsed with Conditions" measures, measure implementers can treat these measures as endorsed, but there are certain conditions, outside of maintenance requirements, that should be met once the measure comes back through for maintenance for the endorsement to be maintained. If a measure with an "Endorsed with Conditions" designation is evaluated for endorsement review, but has not met the prior conditions, then the committee may choose to remove endorsement, unless it agrees with any rationale provided by the devleoper/steward.

Lastly, this E&M Guidebook will continue to evolve, based on the needs of interested parties, the availability of new data sources, the advances in quality measurement science, etc. Before implementing any significant changes to its E&M process, Battelle will conduct informational webinars and a public comment period to solicit input and determine if additional changes are needed. Battelle will continue to work with developers and stewards on their readiness to submit measures to Battelle based on any implemented changes to the E&M process.

Quorum and Voting

Battelle received 17 comments pertaining to quorum and voting. Commenters requested for more information on how an endorsement decision category is reached by using "Met", "Not Met, but Addressable", and "Not Met" ratings for each domain of the PQM Measure Evaluation Rubric? What is the endorsement decision if a measure does not reach consensus? Others expressed the need for developers/stewards to be able to clearly understand the rationale for a committee decision, noting any deficiencies in the submission or the measure specifications or the failure to identify deficiencies in the submission. Some comments drew concern with the 80% voting quorum threshold for a committee of up to 60 people. Lastly, some comments asked how Battelle will ensure committee members who need to vote off-line have the necessary information to make an informed vote and other comments requested that voting should only be done during the endorsement meeting and not offline.

Battelle thanks the commenters for their comments in this area. To address the questions about how and endorsement category is reached, given the ratings of the PQM Measure Evaluation domains, Battelle has provided further clarification, noting the ratings for each domain are used to identify the domains/areas of the measure in which consensus is lacking from the committee, prior to the endorsement meeting. During

the endorsement meeting, Battelle will focus the committee discussions on those non-consensus domains/areas and after facilitated discussion, a final endorsement vote is taken at the end of the meeting. Unlike the prior CBE, there are no "must-pass" domains/criteria in the PQM Measure Evaluation Rubric. If there is still no consensus reached, based on the 75% threshold, the measure is not endorsed. Battelle has also provided more information and examples of how a 75% consensus threshold relates to the 95% index measure of consensus.

Batelle has also noted in the Guidebook that for any endorsement decision, the committee will be asked to provide its rationale, noting any deficiencies in the submission or the measure specifications or the failure to identify deficiencies in the submission.

Battelle appreciates the commenters' concerns with respect to quorum. Battelle's streamlined process will mitigate some of the quorum challenges due to committee member engagement and participation. In addition, Battelle will implement a policy allowing committee members to be inactive for a given cycle and allow Battelle to end a committee-member's term early if there is a consistent lack of participation and engagement (i.e., not attending a meetings without appropriate notice, not submitting independent reviews). However, if voting quorum is not achieved at meeting attendance or is lost during the meeting, Battelle will provide a meeting recording of the discussion to absent committee members. Although, not ideal, this mitigation approach allows for the measure to continue through the E&M process.

Appeals

Ten comments were received pertaining to the Appeals process, raising concern that the Appeals process may prove to be little more than a rehash of prior positions from individuals involved in the original decision in favor of the measure. Some comments disagreed with including Battelle staff and co-chairs of the very committee that voted for or against a measure under appeal review. Some comments noted the removal of the prior CBE post-comment proceedings could potentially lead to the perception of less "consensus development" in the process.

Battelle thanks the commenters for their comments in this area. The Appeals process, including the Appeals Committee, is designed to ensure these meetings can be convened quickly and as needed. The inclusion of Battelle staff and committee chairs will reduce the risk of duplicative or contradictory discussions. Additionally, Battelle clarified in this Guidebook that the Appeals Committee consists of all co-chairs from all five E&M projects, not just those from the committee in which an appeal was received for a measure reviewed by that committee. Lastly, Battelle has added that SMEs may be recruited, as needed, to support the discussions. Battelle will employ the same SME-recruitment approach as outlined previously.

PQM Measure Evaluation Rubric

Finally, Battelle received 16 comments pertaining to various domains of the PQM Measure Evaluation Rubric, including Importance, Scientific Acceptability, Equity, and Use and Usability. Some commenters asked for more resources and supporting materials on how a measure steward/developer can effectively complete the application forms and meet evaluation and endorsement criteria. More clarification is needed as to whether there will be a separate measure evaluation criteria guidebook with more detail and specific algorithms to help the committee evaluate the measure against the Rubric.

Battelle thanks the commenters for their comments in this area and has addressed them below and, where needed, provided edits into the PQM Measure Evaluation

Rubric. Of note, none of the edits made significantly impact the measure requirements/expectations for the Fall 2023 cycle.

Importance

Comments received for the Importance domain recommend rewriting the systematic review requirement to ask for the most up-to-date evidence, rather than a systematic review or other study type. Some asked whether the committees will accept a literature review or is a graded systematic review a requirement in order to pass endorsement? Also, does the business case for a measure need to be as extensive as what is outlined in the CMS Blueprint? Others noted it is unclear what "net benefit" is referring to in relation to the conclusions of the systematic review. There was also disagreement that this domain is "Not Met" due to low certainty of a gap, as measures that may be topped out may still have gaps within subsets of populations.

Battelle has updated the requirement of a systematic review to a literature review, noting various study types can be submitted as evidence to support the measure, including systematic reviews, clinical practice guidelines, observational studies, case studies, etc. The purpose of the literature review is to identify relevant studies and provide a summary of that evidence (which can include various study types) for the committee's consideration. An evaluation of the quality of evidence should also be conducted. Often clinical practices guidelines provide systematic reviews and quality assessment grading of the evidence. If a literature review is not possible, a rationale as to why will be considered by the committee. Battelle has also clarified within the Rubric that the "net benefit" is related to the logic model. Developers/stewards should succinctly summarize the rationale and justification for the measure, supporting this argument with evidence. This aligns with the business case development within the CMS Blueprint. Particularly for new measures, the evidence summary should identify improved outcomes or adverse events and/or costs avoided due to the measure's anticipated impact. Lastly, for performance gap, Battelle has noted within the Rubric that a low certainly of a gap not only applies to overall performance and variation, but also to subsets of populations.

Scientific Acceptability (i.e., Reliability and Validity)

Comments received for the Scientific Acceptability domain requested that until the new SMP gets a chance to provide inputs, the default reliability threshold should be 0.4, as established through Federal rulemaking. There is insufficient information on the criteria to determine whether face validity is "adequate." Some disagreed with the Scientific Acceptability criterion being "Not Met" if results of the risk adjustment model "do not demonstrate acceptable model performance." Commenters shared this creates the possibility for the reviewers to be subjective and inconsistent across measures without a definition of acceptable model performance. Additionally, will developers be required to provide a sufficient level of detail on whether social risk factors were considered and tested? Some comments drew attention to the eCQM testing requirement of at least two electronic health record (EHR) vendors within at least two health system testing is too burdensome. Another comment noted that if developer/steward has challenges with conducting beta testing at a provider/accountable entity-level, this is likely to undermine future use of a potentially impactful measures, such as PRO-PMs and eCQMs, before development is completed. Lastly, one commenter noted the threats to validity, including exclusions and missing data, appear to be omitted from the Rubric.

With respect to the reliability thresholds listed within the PQM Measure Evaluation Rubric, these were confirmed by the SMP during the June 14, 2022 advisory meeting,

after reviewing an impact analysis to determine if endorsement decisions will change substantially with the added thresholds. Therefore, no changes were made to reliability thresholds within the Rubric. For face validity testing, there should be a systematic approach to assessing face validity, which will inform whether the measure score is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance. Battelle has provided this added detail to the Rubric. Regarding the omission of certain threats to validity, Battelle asks about missing data and exclusions as part of the validity testing, specifically,

> **For each level of testing conducted, describe the method of validity testing and what it tests.**
> *Describe the steps, do not just name a method and what you tested (e.g., accuracy of data elements compared with authoritative source, relationship to another measure as expected). What statistical analysis did you use? Include analysis of missing data and any exclusions.*

Regarding the risk adjustment comments, Scientific Acceptability is "Not Met" for a variety of factors. However, if the measure is risk-adjusted, model performance is not the sole reason for determining if validity is "Met" or "Not Met". Rather, the model performance is taken into consideration along with other factors, such as the impact on performance for providers at low and high extremes, in addition to others. When considering what to include in the risk adjustment model, all potential risk factors (clinical, social, and functional) should be considered at the conceptual model phase of the overall risk adjustment approach. If a measure is risk-adjusted, developers/stewards should articulate the rationale for including/not including identified factors. This rationale should be based on the conceptual model and should consider such things as data availability, whether the risk factor is present at the start of care, whether a risk factor will perpetuate inequities in care, the ability of the accountable entity to influence the risk factor, and the factor's influence the measured outcome.

Lastly, Battelle very much appreciates the burden concerns with the eCQM testing and how the Rubric will allow for measure review given challenges with beta testing. Battelle has edited the eCQM testing requirements to state, in addition to other requirement for eCQMs (e.g., feasibility assessment, industry standard specifications), the minimum requirement is testing in at least two EHR vendors, regardless of the number of health systems within which the measure was tested. Additionally, Battelle appreciates and recognizes there are situations in which measures require data element testing prior to implementation for prospective, accountable entity-level testing (e.g., eCQMs and PRO-PMs). The "Endorsed with Conditions" endorsement decision may apply here, in which a new eCQM or PRO-PM has conducted initial testing of the specifications or of the instrument used within a measure, but the developer/steward has not been able to conduct accountable entity-level testing. The E&M committee may determine the measures are endorsed with the condition of evaluating accountable entity-level testing by the time of maintenance review, pending sufficiency across other domains of the Rubric.

Use and Usability

One comment received expressed concern with asking for evidence of behavioral change by the measured entities is making an implicit assumption the measure itself is an intervention. The commenter recommended removing evidence of behavioral change from the use and usability criterion, as the Pre-Rulemaking Recommendations process is more suited for evaluating the policy context.

Battelle would like to clarify there is no requirement or criteria listed within the Use and Usability domain of the PQM Measure Evaluation Rubric asking for evidence of behavior change. Rather, this domain is assessing whether and how the measure is used or will be used. Developers/stewards should provide a summary of the actions accountable entities can take to improve measure performance. For maintenance measures, there is an expectation that the measure is used and developers/stewards should summarize the approach to soliciting feedback about the measure, especially if the feedback led to a change in the measure specifications. In addition, developers/stewards should summarize any progress on improvement. This can be done by looking at trends in performance results over time, including performance among sub-populations, if available, number and percentage of people receiving high-quality health care, geographic area, number and percentage of accountable entities and patients included.

Equity

A few comments received asked whether this criterion is intended to replace the previous sub-criterion on disparities in care. Others raised concern with this new criterion, since not all measures are designed for, nor can identify patient disparities (e.g., structure measures). Therefore, this criterion should be removed or made optional.

Battelle has noted this criterion/domain is optional, as Battelle recognizes some measures are not designed to identify disparities. Battelle is continuing to explore this, but to align with national priorities, Battelle encourages developers/stewards to address this domain, if and when possible.

Additional Supporting Materials/Resources

Battelle recognizes the additional resources and/or guidance may be needed for developers, stewards, and committee members. Namely, Battelle will develop supporting materials, such as a What Good Looks Like" for the Spring 2024 cycle. Unfortunately, this will not be made available for the Fall 2023 cycle. However, Battelle continues to provide technical assistance and flexibility for developers/stewards, where possible.

Regarding the request for an algorithm, currently, the PQM Measure Evaluation Rubric guides reviewers to a decision of "Met", "Not Met, but Addressable", or "Not Met" for each domain, based on set of criteria, replacing the need for an algorithm. However, Battelle will consider if further guidance, such as an algorithm, is needed in the future. In addition, Battelle will host an informational webinar in December 2023 to provide an overview of the PQM Evaluation Rubric and how it is applied.


Partnership for
Quality Measurement
Powered by Battelle