



PROCEDURE MANUAL FOR PERFORMANCE MEASURES DEVELOPMENT AND MAINTENANCE

A Voluntary Consensus Process

JUNE 2016

2016 American Dental Association on behalf of the Dental Quality Alliance (DQA) ©. All rights reserved.
Use by individuals or other entities for purposes consistent with the DQA's mission and that is not for
commercial or other direct revenue generating purposes is permitted without charge

Purpose

This Manual is developed by the **DQA Measures Development and Maintenance Committee** and serves as the basis for developing standardized performance measurement in dentistry. The Manual is updated on a periodic basis as determined by the DQA. For more information on the DQA, please access www.ada.org/dqa or contact dqa@ada.org

Table of Contents

Contents

Overview4

 Objectives4

Roles4

 Measures Development and Maintenance Committee (MDMC)4

 MDMC Chair5

 Conflict of Interest Procedures6

 Confidentiality8

 Copyright Agreement8

Measure Development and Maintenance Process Overview9

 Measure Identification10

 Compiling list of existing measure concepts: Environmental scan.....10

 Initial review of measure concepts and gap identification.....10

 Importance11

 Validity12

 Feasibility13

 Evaluating evidence to support measures14

 Developing draft measure specifications15

 Developing the Concept Report15

 Measure Evaluation16

 Developing a request for proposals16

 Overseeing and guiding feasibility, reliability and validity testing17

Issuing interim report of testing results.....	19
Voting on the fully specified and finalized measures.....	20
Developing the final report	20
Measure Dissemination.....	21
Measure Maintenance Process	21
Appendix 1: Conflict of Interest and Confidentiality Disclosures	23
Appendix 2: Domains of Health Care Delivery Measures.....	28
Appendix 3: Sample Specification for Administrative Measures.....	30
Appendix 4: Request for Proposals Terms.....	31
Appendix 5: Sample Measure Testing RFP Evaluation Summary Sheet (Administrative Data)	33
Appendix 6: NQF Criteria for Endorsement Guidance	34

Overview

The Dental Quality Alliance (DQA) was established to lead efforts in the development of performance measures for oral health care. The DQA is an organization of major stakeholders in oral health care delivery that uses a collaborative approach to develop oral health care measures. The **mission** of the DQA is **to advance performance measurement as a means to improve oral health, patient care and safety through a consensus-building process.**

Objectives

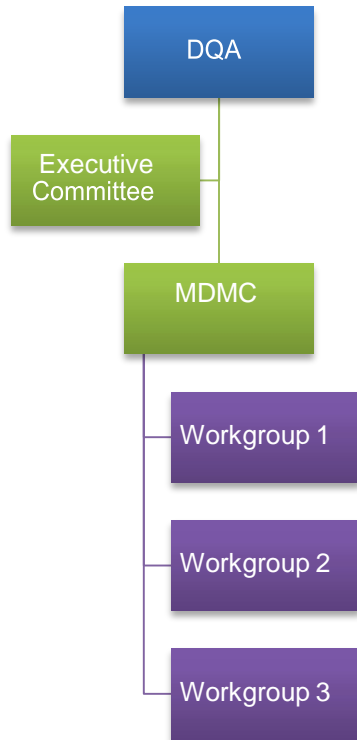
1. To identify and develop evidence-based oral health care performance measures and measurement resources.
2. To advance the effectiveness and scientific basis of clinical performance measurement and improvement.
3. To foster and support professional accountability, transparency, and value in oral health care through the development, implementation, and evaluation of performance measurement.

Performance measures are developed through a consensus process based on the best available evidence. The process also identifies gaps in measures and limitations of the current data infrastructure. This procedure manual documents how the DQA develops and maintains measures through a process that is collaborative, objective, transparent, and meaningful.

Roles

Measures Development and Maintenance Committee (MDMC)

The Measures Development and Maintenance Committee (MDMC) of the DQA oversees measure development and maintenance. The measure development and testing process entails initial selection of oral health care topic areas by the DQA. The MDMC refines the topic areas and oversees ad hoc workgroups that identify measure concepts and develop detailed measure specifications. In addition, the MDMC oversees the measure maintenance processes. The measure maintenance process includes annual review of the measures and the User Guide. The DQA approves the final work products of the MDMC and its workgroups.



The DQA strives to ensure that the measure development and maintenance process remains objective, transparent, and collaborative. To this end, all organizations within the DQA have multiple opportunities to review and provide input during the measure development and maintenance process.

DQA members nominate subject matter experts to the MDMC and the workgroups. Subject matter experts should be (1) currently active and respected in their field; (2) capable of knowledgeably participating in the measure development activities; and

(3) available to participate in conference calls and face to face meetings. These individuals do not represent any organization but rather serve as individuals/subject matter experts on the MDMC and its workgroups. Documents published by the DQA acknowledge the contribution of these individuals.

MDMC Chair

The DQA Chair designates the MDMC Chair from among selected nominees. The selection of the Chair is based on candidate's experience in developing quality measures and absence of any significant conflicts of interest with the project. The Chair should be skilled in chairing meetings, possess basic knowledge of parliamentary procedure and the proper role of the chair as a neutral facilitator, be skilled in scientific writing, have prior experience in leading expert discussions, and be capable of facilitating the interpersonal aspects of group processes so that the panelists work in the spirit of collaboration with balanced contribution from all members. The Chair should be capable of meeting the following commitments:

- Understand the process for developing and maintaining measures as described in this manual;
- Assist staff in planning meeting agendas;
- Moderate and guide the Committee during its development and maintenance of measures;
- Ensure that the group functions effectively and remains focused;
- Encourage all members of the group to contribute to the discussions;
- Delegate assignments and integrate completed assignments and group feedback into draft report;
- Stimulate discussion and facilitate group consensus while refraining from undue personal input; and
- Encourage constructive debate without forcing agreement.

Conflict of Interest Procedures

To ensure that a collaborative and balanced approach is followed, the DQA requests that all individuals nominated to the MDMC and its workgroups complete a standard conflict of interest form ([Appendix 1](#)).

Disclosed conflicts are not confidential. Unless the individual is disqualified to serve, his or her disclosures will be shared with the other members and published with the report. Disclosure allows the DQA to maintain a transparent process and convene a balanced group.

The DQA Chair and Chair-Elect will review disclosures of nominees and determine each nominee's eligibility to serve and/or vote on the final recommendations. Completed disclosure forms will be kept on file by DQA staff. Each nominee will be notified by DQA staff of the determination by the Chair and Chair-Elect.

Individuals may recuse themselves voluntarily from participation with regard to specific aspects of the processes; however, a voluntary recusal does not free a member from the obligation to disclose a conflict.

All persons who develop potential conflicts of interest after initial disclosure must update the Conflict of Interest Questionnaire and disclose changes by electronic means to the DQA Chair.

Procedures for review of completed disclosure forms and rules for action

The DQA Chair's and Chair-Elect's ruling on the person's eligibility to participate and/or vote on the Committee/Workgroup will consider the following:

- Is there any question that the person has not made a full and complete disclosure?
- Is there any indication that the person may provide any information that could be perceived as misleading?
- Is there any indication that the person while participating in the Committee/Workgroup may improperly favor any outside entity or may appear to have an incentive to do so?
- Does the person appear to be subject to incentives that might lead to disqualifying bias?
- Is there any indication that the person's conflict may prevent him or her to meet his or her obligations to, or the objectives of, the designated project?
- Do the person's current engagements present any conflicts between outside interests (e.g., is he simultaneously working on projects for competing business entities, fiduciary positions with other organizations, etc.)?

The DQA Chair and Chair-Elect will make a determination of appropriate action. The following rules will apply.

- *No action.*

No disclosure or recusal necessary and individual may fully participate in the Committee/Workgroup's activities

- *Information disclosure to Committee/Workgroup.*

Individual must disclose potential conflict to the full Committee/Workgroup and may fully participate in discussion and vote.

- *Information disclosure to Committee/Workgroup and recusal from voting.*
Individual must disclose potential conflict to the Committee/Workgroup and may fully participate in discussion but will be recused from voting.
- *Disqualification from all participation*
Individual may not be part of the Committee/Workgroup.

Procedures for voting

At the discretion of the MDMC Chair, votes may be taken for major procedural and methodological decisions during the measure development process. Voting procedures include the following:

- Votes are taken by voice or hand, without secret ballots.
- A quorum for official votes is at least one-half of eligible members (those not specifically recused for disclosed conflicts), including the chair.
- Reconsideration of a previously voted statement requires approval of two-thirds of those eligible to vote.
- Ex-officio members do not vote.

Confidentiality

All discussions and documents should remain confidential **until the interim and final reports are publicly disseminated**. If workgroup members are provided access to embargoed publications during the course of the discussions, such information should remain confidential until final publication. ([Appendix 1](#))

Copyright Agreement

All DQA Volunteers are required to sign a copyright agreement such that intellectual property right for the materials developed during DQA Committee/Workgroup work is appropriately transferred to the DQA. ([Appendix 1](#))

Measure Development and Maintenance Process Overview

Measure Development Process

The process of developing measures typically occurs in three phases.

Phase 1: Measure Identification

- Compiling a list of existing measures: Environmental scan
- Initial review of existing measure concepts and identification of measurement gaps
- Evaluating evidence to support measures
- Developing draft measure specifications
- Developing the Concept Report

Phase 2: Measure Evaluation

- Developing a Request for Proposals/Statement of Work for measure testing
- Overseeing and guiding feasibility, reliability and validity testing
- Issuing Interim Report of testing results
- Voting on fully specified and finalized measures
- Developing the Final Report

Phase 3: Measure Dissemination

Measures Maintenance Process

The measures developed by the DQA undergo periodic review to assess impact and potential unintended consequences.

Phase 1: Call for Comments

- Release call for comments to the measures and the User Guide

Phase 2: Review & Evaluation

- Review submitted comments and proposed changes
- Conduct additional testing as needed
- Issue Draft Report that includes all proposed changes to the measure specifications and the User Guide

Phase 3: Approval

- Voting by the DQA on proposed changes
- Develop Final Report, with revised specifications and User Guide
- Approved changes published to the DQA website with effective date of January 1st the following year.

The ensuing sections of this manual describe these steps in more detail.

Measure Identification

Compiling list of existing measure concepts: Environmental scan

The Committee/Workgroup begins its work by identifying existing performance and quality measure concepts on the assigned topic. A comprehensive scan is available from the initial work of the MDMC in 2012. This scan is updated as needed.

Environmental scan resources include:

1. PubMed searches
2. Keyword searches of the internet using standard search engines such as google
3. Searches through the links provided within the National Library of Medicine database of relevant organizations (<http://www.nlm.nih.gov/hsrinfo/quality.html#760>) including the National Quality Measures Clearinghouse (NQMC), National Quality Forum (NQF), Maternal and Child Health Bureau (MCHB), etc.
4. Soliciting measures from other measure development organizations (e.g., Veterans Administration, public and private payers, HRSA programs)
5. Other sources identified by Committee/Workgroup members

Initial review of measure concepts and gap identification

The goal for the initial review of concepts is to identify existing concepts that are important, valid, and feasible. Data for measurement in dentistry is obtained from administrative sources (claims and encounters), patient records within electronic systems (e.g., Practice Management Software and EHR systems), and patient surveys. The construct of measures is affected by the data available from each of these sources. Thus, feasibility depends on the data source that will be used for implementation (i.e., administrative claims vs. dental records/EHR vs. surveys).

Note that the rating of concepts at this stage in the process is based solely on the knowledge and expert judgment of the Committee/Workgroup members. Once an initial

set of measure concepts is identified, they are presented to the DQA for approval for testing and further development. In instances when the Committee/Workgroup is faced with a large set of measure concepts for review, the Chair may choose to use a Delphi process to facilitate consensus using the RAND-UCLA modified Delphi approach.¹ Criteria for this rating exercise shall be based on those used by the Agency for Healthcare Research and Quality (AHRQ).²

The AHRQ process also provides quantitative scoring criteria that may be used to winnow down the list of concepts. Alternatively, the Committee/Workgroup can pare down the list first based on importance of the concept. For those concepts deemed important, feasibility and validity may then be assessed. Concepts that are deemed important and valid but not feasible may be used to provide recommendations for structured data elements that may be necessary to support future quality measures. In cases where the environmental scan results in a manageable number of measure concepts, the Chair may request the Committee/Workgroup to discuss each measure individually.

The AHRQ process identifies the following considerations when evaluating measure concept **importance**, **validity**, and **feasibility**.³

Importance

To be considered important at least some of the following criteria should be met by the measure.

1. The measure should be actionable. States, Medicaid and CHIP managed care plans, and relevant health care organizations should have the ability to improve their performance on the measure with implementation of quality improvement efforts;
2. The cost to the nation for the area of care addressed by the measure should be substantial;

¹ Brook RH. The RAND/UCLA appropriateness method. In: McCormick KA, Moore SR, Siegel RA eds. Clinical Practice Guidelines Development. Methodology Perspectives. Rockville, MD: Agency for Health Care Policy and Research; 1994

² Mangione-Smith R, Schiff J, Dougherty D. Identifying children's health care quality measures for Medicaid and CHIP: an evidence-informed, publicly transparent expert process. Acad Pediatr. 2011 May-Jun;11(3 Suppl):S11-21.

³ The importance, validity, and feasibility criteria described here follow Mangione-Smith R, Schiff J, Dougherty D. Identifying children's health care quality measures for Medicaid and CHIP: an evidence-informed, publicly transparent expert process. Acad Pediatr. 2011 May-Jun;11(3 Suppl):S11-21.

3. Health care systems should clearly be accountable for the quality problem assessed by the measure;
4. The extent of the quality problem addressed by the measure should be substantial;
5. There should be documented variation in performance on the measure;
6. The measure should assess an aspect of health care where there are known disparities.

Validity

Validity is the degree to which a quality measure is associated with what it purports to measure (e.g., a clinical decision support system is a measure of structure or capacity; prescribing is a measure of a clinical process; asthma exacerbations are a measure of health outcomes).

A quality measure should be considered *valid* if:

1. There is adequate scientific evidence or, where evidence is insufficient, expert professional consensus to support the stated relationship between:
 - **structure and process:**⁴ e.g., that there is a demonstrated likelihood that a clinical decision support system (a structural or capacity measure) in a hospital or ambulatory office leads to increased rates of appropriate flu vaccination in the hospital or practice,
 - **structure and outcome:** e.g., higher continuity of care in the outpatient setting (influenced by how appointments are organized) is associated with fewer ambulatory care sensitive hospitalizations (e.g., hospitalizations for dehydration), or
 - **process and outcome:** e.g., that there is a demonstrated likelihood that prescribing inhaled corticosteroids (a clinical process) to specified patients with asthma will improve the patients' outcomes and vice versa (e.g., that if

⁴ Structure of care is a feature of a healthcare organization or clinician relevant to its capacity to provide health care. A process of care is a health care service provided to, on behalf of, or by a patient appropriately based on scientific evidence of efficacy or effectiveness. An outcome of care is a health state of a person resulting from health care. National Quality Measures Clearinghouse: www.qualitymeasures.ahrq.gov.

we measure quality as a health outcome measure there is sufficient demonstrated likelihood that the outcome can be attributed to either health care delivery structures or clinical processes of care or an explicit combination of both).

2. The health care system can be said to be responsible for performance and/or the related health outcome. The majority of factors that determine adherence to a measure are under the control of the clinician, clinic, hospital, health plan, or the Medicaid or CHIP program subject to measurement.

Feasibility

A quality measure will be considered feasible if:

1. The information necessary to determine adherence to the measure is likely to be found in available data sources (e.g., administrative billing data, structured data in electronic records, or routinely collected survey data).
2. Estimates of adherence to the measure based on available data sources are likely to be reliable and unbiased. Reliability is the degree to which the measure is free from random error.

Following the rating process, the Committee/Workgroup may find:

1. Measure concepts that are complete and have complete measure specifications: The Committee/Workgroup shall acknowledge such measures and provide as much detail in their report with links to the source/organization that developed the measure.
2. Measure concepts that are complete as written but do not have complete measure specifications: If the Committee/Workgroup believes that the concepts are complete but they lack accompanying specifications, the Committee/Workgroup shall contact the source of the concept and collaborate to fully specify the concept.
3. Measure concepts that express a theme but are found to be lacking in detail and do not have specifications: The Committee/Workgroup shall develop *de novo* measure concepts and specifications based on these themes.
4. Other aspects of health care that do not have existing concepts: If the

Committee/Workgroup believes that there are other guidelines that address important issues and do not have applicable measures, they should develop *de novo* measure concepts.

Evaluating evidence to support measures

Once all relevant concepts are identified for the assigned topics, the Committee/Workgroup categorizes the measures based on the domain framework developed by the National Quality Measure Clearing House (NQMC) ([Appendix 2](#)). As noted earlier, to be classified as a clinical quality measure, the following considerations apply as recommended by NQMC:⁵

- For process measures, evidence that the measured clinical process has led to improved health outcomes.
- For outcome measures, evidence that the outcome measure has been used to detect the impact of one or more clinical interventions.
- For access measures, evidence that an association exists between the access measure and the outcomes of or satisfaction with care.
- For patient experience measures, evidence that an association exists between the measure of patient experience of health care and the values and preferences of patients/consumers.
- For structure measures, evidence that an association exists between the structure measure and one of the other four domains of quality listed above (e.g., process, outcome, access, and patient experience).

The Committee/Workgroup should document at least one of the following types of evidence within its final report:

- a clinical practice guideline/recommendations or other peer-reviewed synthesis of the clinical evidence,
- a systematic review of the clinical literature, and/or
- one or more research studies published in a National Library of Medicine (NLM)

⁵ National Quality Measures Clearinghouse Domain Definitions:

<http://www.qualitymeasures.ahrq.gov/about/domain-definitions.aspx>. Accessed February 2013.

indexed, peer-reviewed journal.

Additional guidance on evidence is available from the [National Quality Forum](#).

Developing draft measure specifications

The Committee/Workgroup then defines preliminary measure specifications for each concept. A template for the measure specification for measures based on administrative data is available in [Appendix 3](#).

Preliminary specifications must include as much detail on the measure logic and the codes as possible with specific notations on what information is missing. The more detailed the specifications at this stage, the easier it is for the dental community to assess the measure and provide feedback to determine consensus.

An important focus of measurement to improve quality of care is the study of variations (by age, race/ethnicity, socioeconomic status, length of enrollment, geographic area, plan type, etc.) in care. Use of such stratification variables provides an important tool to understand variations in care. Appropriate stratification variables should be identified for each measure.

Developing the Concept Report

The Committee/Workgroup develops an interim report with the list of proposed measures. When developed by a Workgroup, the MDMC must approve the proposed measures.

Proposed measures are then routed to the broader DQA for comment. Based on the evaluation of the comments, the MDMC makes a recommendation to the DQA about whether to move forward with measure testing.

Measure Evaluation

For proposed measures that have consensus for testing, the MDMC proceeds with establishing measure feasibility, validity, reliability, and usability.

Developing a request for proposals

The DQA may use a competitive request for proposals (RFP) process to identify investigators to conduct testing of measures or request a Statement of Work from known entities/DQA members willing to conduct the testing.

When an RFP is issued, it must identify the following:

1. Application Deadline
2. Project Deliverables
3. Minimum and Desired Requirements
4. Guidelines for Information to be Included within the Proposals
5. Evaluation Criteria
6. Terms ([Appendix 4](#))
7. Draft Specifications
8. Guidance for Testing

When an RFP process is used to contract for measure testing, the DQA Chair will appoint a Review Panel to review proposals. Procedures for addressing disclosed conflicts and rules of action are the same as defined earlier in the document. The following are examples of significant conflicts for this stage of the process.

An actual or potential conflict of interest shall be deemed to exist when a potential reviewer:

- a) is the Program Director/Principal Investigator (PD/PI) or one of multiple PDs/PIs;
- b) is a Senior/Key Personnel, other significant contributor, collaborator, or

- consultant;⁶
- c) is a member of an advisory board or research team for the proposal;
 - d) within the preceding three years, has collaborated with, co-authored a publication(s) with, and/or mentored or trained the PD/PI, one of multiple PDs/Pis, or an individual named on the application as participating with a major professional role;
 - e) is in collaboration, is negotiating collaboration, or is preparing an application(s) or publication(s) with the PD/PI, with one of multiple PDs/Pis, or with an individual named in the application as participating with a major professional role for a competing endeavor;
 - f) has written a letter of general support or enthusiasm for the application in question but plays no substantive role in the proposed work; or
 - g) belongs to the organization or entity applying for the program.

Staff will compile and distribute all proposals to the Review Panel members. Specific Panel members may be assigned as leads on specific proposals to manage the workload. A consensus process is used to determine the best proposal that meets the needs of the DQA. Guidance on available funding will be provided by the MDMC in consultation with the DQA Chair. Panel members must use standardized worksheets to evaluate all proposals in an unbiased manner. A sample worksheet for review of proposals based on administrative data is available in [Appendix 5](#). All applicants must be provided with a summary evaluation sheet that lists the strengths, weaknesses, and outcome of the review of their proposals at the end of the process.

Overseeing and guiding feasibility, reliability and validity testing

The MDMC provides oversight and guidance during the measure testing phase. Measures developed by the DQA may be submitted for endorsement by the National Quality Forum (NQF). The NQF requires data for topic importance, performance gap, evidence to support process measures, scientific soundness, feasibility, and use and

⁶ A consultant or collaborator who has received or could receive a direct financial benefit of any amount from an application under review, applicant institution, or PD/PI, or has received or could receive a financial benefit from the applicant institution or PD/PI that in the aggregate exceeds \$10,000/year is defined as a major professional role.

usability. In order to meet the criteria for scientific soundness (reliability and validity), acceptable protocols for testing should be designed to address the NQF evaluation criteria in place at the time testing commences. Assessing usability is to assure that the information produced by the measure is meaningful, understandable, and useful to the intended audience. Complete NQF recommendations are available at the NQF website.⁷ A snapshot of the NQF evaluation criteria is available in [Appendix 6](#).

The final report from the testing effort should provide data to answer the following⁸:

Feasibility

1. To what extent are the data elements necessary to define numerator/denominator and exclusions readily available within one or more databases?
2. Are there certain data elements required to compute the numerator/denominator that are more prone to be incomplete or missing (e.g., claims/encounters and eligibility/enrollment files)?
3. Are there any significant barriers encountered during data collection and measure computation?
4. What were the resources required to calculate this measure set? (personnel and system resources)
5. Were any significant problems encountered due to vague measure definitions and/or specifications?
6. Can an automated report be generated?
7. Is the data element accurate i.e. is it generally captured by the most appropriate person involved in the clinical workflow?
8. For eMeasures, are the data elements necessary to define numerator/denominator and exclusions readily available in a structured format across EHR systems?
9. For eMeasures, to what extent does capturing the data element fit the typical workflow for that user/system?

Reliability

10. Are the results from the measure repeatable?
11. For each measure, have all the data elements required to compute the numerator/denominator and exclusions been identified within the technical specifications?
12. Is the data element coded using a nationally accepted terminology standards?
13. To what extent do the exclusions due to missing or invalid data impact the measurement score? (The National Quality Forum provides additional guidance on testing for threats to validity from missing or "incorrect" data or exclusions (selection/attrition bias) (<http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=59116>) Sensitivity analyses with and without the exclusion, and variability of exclusions across measured entities can be used to determine the impact of missing or incorrect data on the resulting measure.

⁷ Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement. April 2015. Accessed at: https://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx

⁸ National Quality Measures Clearinghouse. Tutorials on Quality Measures: <http://www.qualitymeasures.ahrq.gov/tutorial/index.aspx>; National Quality Forum. Measure Evaluation Criteria: http://www.qualityforum.org/Show_Content.aspx?id=322.

Validity

14. To what extent does the measurement score truly represent what it is intended to measure (compare with published literature)?
15. To what extent is the health care construct underlying the measure associated with important health care processes and/or outcomes (e.g., published literature presents strong evidence for an association).
16. Is there an opportunity for improvement?
17. Are all individuals in the denominator equally eligible for inclusion in the numerator? (A valid measure of quality of care should exclude individuals who should not receive the indicated care or are not at risk for the outcome.)
18. Is the measure result under control of those whom the measure evaluates? (Example: A measure of asthma prevalence within a Health Plan is a not a measure of Outcome but of User/Enrollee Health Status. Clinicians can diagnose asthma, but asthma is primarily caused by genetic and environmental risk factors, not by receiving health care. A user should not use this measure to compare health care providers who care for populations that differ in their risk for developing asthma.)
19. How well do the measure specifications capture the event that is the subject of the measure?
20. For accountability measures, does the measure provide for fair comparisons of the performance of providers, facilities, health plans, or geographic areas? (stratified or risk adjusted)
21. For accountability measures, does the measure allow for adjustment of the measure to exclude patients with rare performance-related characteristics when appropriate? (A measure concerning provision of an evidence-based treatment allows exclusion of patients who refuse the treatment.)
22. For accountability measures, are the measure thresholds or targets appropriately identified?

Usability

23. To what extent are the measure rationale and results easily understood by users of the measure and resulting data?
24. To what extent are there performance gaps or significant variation among measured entities that can be addressed by implementing the measure?
25. To what extent are the measure results reportable in manner useful to health care organizations and other interested stakeholders?

The MDMC oversees and works closely with the testing team to iteratively finalize the measure specifications as the testing progresses.

Issuing interim report of testing results

After the majority of testing is completed, an interim report is prepared with input from the testing team, DQA staff, and the MDMC. The report summarizes the data sources, testing methodology and results to date along with key determinations made by the MDMC. Updated measure specifications are included in the report. Once the report is approved by the MDMC, it is released for a 30-day public comment period.

Dissemination methods include electronic communication to key stakeholders and posting the report online. Each comment is reviewed and addressed by the MDMC with additional testing and refinement of the measure specifications as needed.

Voting on the fully specified and finalized measures

After testing is completed, the MDMC votes on whether to recommend the measure for approval to the full DQA. A draft final report is prepared that includes the key testing results and findings of the MDMC. The report, recommendations, and finalized measure specifications are presented to the full DQA for consideration. Following the criteria for measure importance, feasibility, reliability, validity, and usability, the DQA votes on whether to approve the measure.

Developing the final report

After the DQA has voted on whether to approve a measure, a final report is prepared that is a comprehensive document that details the data sources, testing methodology, testing results, and finalized measure specifications. The rationale and supporting data for key determinations made during testing should be documented. The report should address the evaluation criteria of importance, feasibility, reliability, and usability. For approved measures, the report should provide the requisite details to support submission of the measures to the National Quality Forum (NQF) and other endorsement agencies. This report is typically built on the interim report and may be organized by the following subhead titles:

- Abstract
- Scope and purpose
- Data sources
- Testing methodology
- Evidence for validity
- Evidence for reliability
- Evidence for feasibility
- Evidence for usability (performance gap)
- Final measure specifications along with calculation algorithms

- Defined sampling procedures (if applicable)
- Risk adjustment (if needed)
- Implementation considerations (including potential obstacles to implementation)

Measure Dissemination

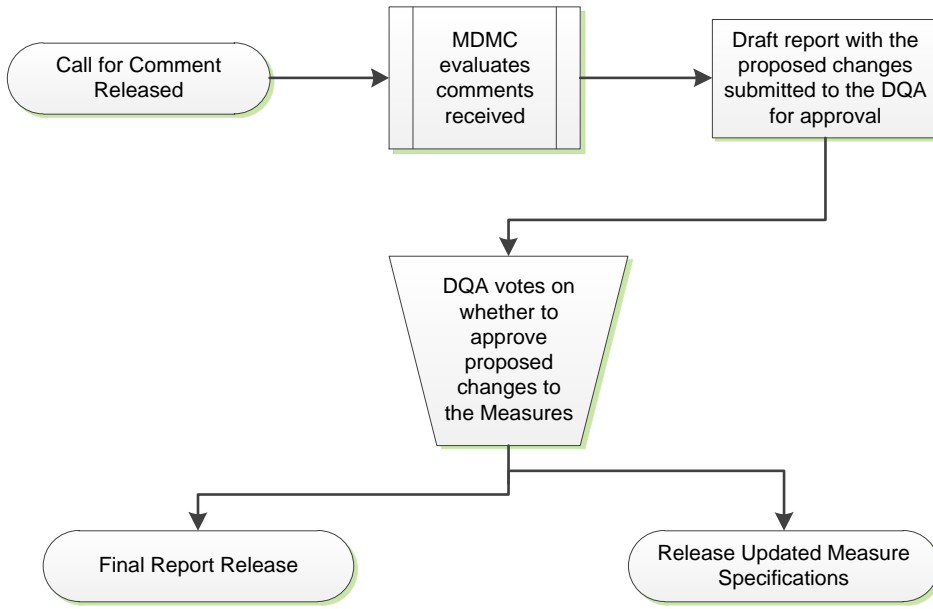
Several modes may be considered for disseminating measures:

- Posting on DQA website
- Peer-reviewed journal publications – articles and a one-page executive summary
- Submission to the National Quality Measures Clearinghouse
- Submission to the National Quality Forum
- Communication through DQA member e-communications
- Conference presentations
- Webinars

Available resources are directed to maximize reach to target audiences.

Measure Maintenance Process

In order to ensure transparency and establish proper protocols for timely assessment of the evidence and the properties of the measures, as well as to comply with the NQF's endorsement agreement, the DQA has established a measure maintenance process. The process of the annual measure review follows an annual cycle as depicted in the following flow chart:



Process Timeline: DQA annual measure review follows a cyclical timeline

DQA Annual Measure Review Process Timeline	
Date	Tasks/ Events
February, 20XX	A call for comments is announced with a 30-day comment period
March 1 st - April 30 th , 20XX	MDMC evaluates all comments received; conducts any data analysis that may be required
May, 20XX	Draft report summarizing the evaluation results and any proposed changes to the measure specifications and the User Guide is developed for DQA's review
June/July, 20XX	Proposed changes to the measure specifications and the User Guide voted on by the DQA
September, 20XX	<ul style="list-style-type: none"> Final Report of the Annual Measure Review Released Updated measure specifications released/disseminated via DQA website Updated User Guide released/disseminated via DQA website
January 1 st , 20XX	The updated versions effective January 1 st of the following year.

For more information, please access the DQA website at www.ada.org/dqa or contact the DQA by email at dqa@ada.org.

Appendix 1: Conflict of Interest and Confidentiality Disclosures

Objective

This Conflict of Interest procedure supports the goal of having a process by which the Dental Quality Alliance (DQA) reviews proposals for measure testing that remains consistent, objective, and transparent. All stakeholders must have confidence in the integrity of the process in order to accept the recommendations of the reviewers in identifying suitable investigators to support funding for measure testing.

CONFLICT OF INTEREST

General Procedures

In order to effectively identify conflicts of interest, individuals must disclose any potential conflicts of interest upon being invited to participate in the panel through the Conflict of Interest Questionnaire. The intent and purpose of this disclosure is to avoid total disqualification and to give more guidance to individuals who complete the Questionnaire. Thus, answering "yes" to many or even most of the questions will not lead to disqualification of the individual. Indeed, in many instances it is important to have individuals who have a certain level of expertise which can only be attained by affiliations with other individuals, organizations, or companies.

- The Chair and Chair-Elect of the DQA shall determine the person's eligibility to participate and/or vote on the panel.
- Each person will be notified of the DQA Chairs' ruling by Staff.
- Individuals may recuse themselves voluntarily from participation with regard to specific aspects of the processes; however, a voluntary recusal does not free a member from the obligation to disclose a conflict.
- Completed disclosure forms will be kept on file by staff.
- All persons who develop potential conflicts of interest after initial disclosure must update the Conflict of Interest Questionnaire and disclose changes by electronic means to the Chair of the DQA with a copy to staff.
- Disclosed conflicts will be reported along with the guidance provided to the

individual when final recommendations of the review panel are submitted to the DQA.

Procedures for review of completed disclosure forms and rules for action

The DQA Chairs' ruling on the person's eligibility to participate and/or vote on the panel will consider the following:

- Is there any question that the person has not made a full and complete disclosure?
- Is there any indication that the person has provided any information that could be perceived as misleading?
- Is there any indication that the person while participating in the review panel may improperly favor any entity or may appear to have an incentive to do so?
- Does the person appear to be subject to incentives that might lead to disqualifying bias?
- Is there any indication that the person's conflict may prevent him or her to meet his or her obligations to, or the objectives of, the Review panel?
- Do the person's current engagements present any conflicts between outside interests (e.g., working on projects simultaneously for competing business entities, fiduciary positions with other organizations, etc)?

A determination of appropriate action will be a made by DQA Chair and Chair-Elect.

The following rules will apply.

- *No action.*
 - No disclosure or recusal necessary and individual may fully participate in the panel's activities
- *Information disclosure to expert panel.*
 - Individual must disclose potential conflict to the full panel and may fully participate in discussion and vote.
- *Information disclosure to expert panel and recusal from voting.*
 - Individual must disclose potential conflict to the full panel and may fully participate in discussion but will be recused from voting.

- *Disqualification from all participation*
 - Individual may not be part of the expert panel.

Procedures for Voting

At the discretion of the Workgroup Chair, votes may be taken for major procedural and methodological decisions during the measure development process. Voting procedures include the following:

- Votes are taken by voice or hand, without secret ballots.
- A quorum for official votes is at least one-half of eligible members (those not specifically recused for disclosed conflicts), including the chair of the review panel.
- Reconsideration of a previously voted statement requires approval of two-thirds of those present.
- Ex-officio members do not vote.

Certification

I certify that I have read and understand the description of conflict of interest above and

I do not have any actual or perceived conflicts of interest

OR

I have the following actual or potential conflict of interest. (Please list below)

I have read the DQA Conflict of Interest Policy and understand that I have a continuing responsibility to comply with such policy. I further understand that I am required to promptly disclose any conflict of interest that might arise, as well as any material changes to the answers I have provided in this Conflict of Interest Statement. The facts set forth herein are true and accurate to the best of my knowledge

Reviewer's printed name:

Reviewer's signature:

Date:

CONFIDENTIALITY

All discussions and documents related to the measure development process should remain confidential.

Certification

I fully understand the confidential nature of the measure development process and agree: (1) to destroy or return all materials related to the process; (2) not to disclose or discuss the materials associated with the process, my discussions, or the meetings outside of that meeting or with any other individual except DQA staff and members of the Workgroup; and (3) to refer all inquiries concerning the review DQA Chair.

Reviewer's printed name:

Reviewer's signature:

Date:

COPYRIGHT ASSIGNMENT AGREEMENT

The undersigned is participating as a Volunteer on the Dental Quality Alliance (DQA) and/or in Committees and Workgroups. In this capacity, the undersigned's responsibilities for the DQA may include creating, or contributing to the creation of, original content for one or more of the DQA's ongoing publications or for a special project that may result in a publication distributed by the DQA.

The undersigned irrevocably grants, assigns, and transfers to the DQA all right, title, and interest including, but not limited to, any and all copyrights and other intellectual property rights, in and to any original, copyrightable material ("materials") created by the undersigned in his or her capacity as a Volunteer. In addition, to the extent that any such material is covered by one or more of the definitions contained in the United States Copyright Act ("Act"), specifically in 17 U.S.C. & 101, and to the extent all other requirements pertaining to "works made for hire" are satisfied, the

undersigned agrees that such materials may be treated by the Dental Quality Alliance as “works made for hire”. The undersigned understand that he or she is acting as an independent contractor respecting volunteer work performed for the DQA, and shall have no copyright or other right, title, or interest in and to the material, or to any derivative works based thereon, all such material and derivative works being the DQA's sole property.

The undersigned represents and warrants that: (1) he or she has a full power and authority to enter into this Agreement and to grant all rights, interests, and title as provided herein; and (2) he or she will execute any additional documents necessary to give this Agreement full force and effect.

Signature

Name (Please print or type)

Witness

Date

Appendix 2: Domains of Health Care Delivery Measures

The following table presents the definitions of the quality domains as defined by NQMC.⁹ Measures of care delivered to individuals and populations defined by their relationship to clinicians, clinical delivery teams, delivery organizations, or health insurance plans.

Denominators for these measures are defined by some form of affiliation with a clinical care delivery organization, e.g. recipients of health care, health plan enrollees, clinical episodes, clinicians, or clinical delivery organizations.

<p>Clinical Quality Measures: Measures used to assess the performance of individual clinicians, clinical delivery teams, delivery organizations, or health insurance plans in the provision of care to their patients or enrollees, which are supported by evidence demonstrating that they indicate better or worse care</p>	
<p>Process</p>	<p>A process of care is a health care-related activity performed for, on behalf of, or by a patient. Process measures should be supported by evidence that the clinical process that is the subject of the measure has led to improved outcomes. These measures generally are calculated using patients eligible for a particular service in the denominator, and the patients that either do or do not receive it in the numerator.</p> <p>Example: Percentage of enrolled children in the age category of 6–9 years at “elevated” risk (i.e., “moderate” or “high”) who received a sealant on a permanent first molar tooth within the reporting year.</p>
<p>Access</p>	<p>Access to care is the attainment of timely and appropriate health care by patients or enrollees of a health care organization or clinician. Access measures should be supported by evidence that an association exists between the measure and the outcomes of or satisfaction with care.</p> <p>Example: Percentage of all enrolled children who received an oral evaluation within the reporting year</p>
<p>Outcome</p>	<p>An outcome of care is a health state of a patient resulting from health care. Outcome measures should be supported by evidence that the measure has been used to detect the impact of one or more clinical interventions. Measures in this domain should be attributable to antecedent health care and should include provisions for risk-adjustment.</p> <p>Example: Number of emergency department visits for caries-related reasons per 100,000 member months for all enrolled children.</p>

⁹ National Quality Measures Clearing House Domain Framework and Inclusion Criteria. Accessed at <http://www.qualitymeasures.ahrq.gov/about/domain-definitions.aspx>

Structure	<p>Structure of care is a feature of a health care organization or clinician related to its capacity to provide high quality health care. Structure measures should be supported by evidence that an association exists between the measure and one of the other clinical quality measure domains.</p> <p>These measures can focus on either health care organizations or individual clinicians.</p> <p>Example: Percentage of dental offices equipped with autoclaves to ensure proper infection control.</p>
Patient Experience	<p>Experience of care is a patients or enrollee's report concerning observations of and participation in health care.</p> <ul style="list-style-type: none"> • Patient experience measures should be supported by evidence that an association exists between the measure and patients' values and preferences, or one of the other clinical quality domains. • These measures may consist of rates or mean scores from patient surveys. <p>Example: Percentage of enrollees reporting unmet dental care needs.</p>
<p>Related Health Care Delivery Measures: Measures used to assess the non-quality aspects of performance of individual clinicians, clinical delivery teams, delivery organizations, or health insurance plans in the provision of care to their patients or enrollees. These measures are not supported by evidence demonstrating that they indicate better or worse care.</p>	
Use enrollee health state	<p>User-Enrollee Health State: A user-enrollee health state is the health status of a group of persons identified by enrollment in a health plan or through use of clinical services.</p> <p>By definition, a user-enrollee health state is not known to be the result of antecedent health care.</p> <p>Example: Prevalence of complete tooth loss (edentulism) among dental plan enrollees (inclusion in the denominator is based on membership in a particular health plan; however, the measured health state is not a result of that membership). Note: A measure of edentulism prevalence within a Health Plan is a not a measure of Outcome but of User/Enrollee Health Status.</p>
Management	<p>Management of care is a feature of a health care organization related to the administration and oversight of facilities, organizations, teams, professionals, and staff that deliver health services to individuals or populations.</p> <p>Management measures assess administrative activities important to health care but not part of the direct interaction between individual patients and health care professionals.</p> <p>Examples: Whether a practice has a policy to ensure the prevention of fraud and has defined levels of financial responsibility and accountability for staff undertaking financial transactions.</p>
Use of services	<p>Use of services is the provision of a service to, on behalf of, or by a group of persons identified by enrollment in a health plan or through use of clinical services.</p> <p>Use of service measures can assess encounters, tests or interventions that are not supported by evidence of the appropriateness of the service for the specified individuals.</p> <p>Example: Percentage of enrollees who received a dental treatment service.</p>

<p>Cost</p>	<p>Costs of care are the monetary or resource units expended by a health care organization or clinician to deliver health care to individuals or populations. Cost measures are computed from data in monetary or resource units. Costs may be reported directly (i.e. actual costs) or estimated based on the volume of resource units provided and the charges for those units.</p> <p>Example: Total amount that is paid on direct provision of care (reimbursed for clinical services) per member per month for all enrolled children during the reporting year.</p>
<p>Clinical Efficiency measures: Measures that may be used to assess efficiency directly (e.g., by comparing a measure of quality to a measure of resource use) or indirectly (e.g., by measuring the frequency with which health care processes are implemented that have been demonstrated by evidence to be efficient).</p>	
<p>Efficiency of care</p>	<p>Efficiency of care is the propensity of a health care organization or clinician to maximize the number of comparable units of health care delivered for a given unit of health resources used.</p> <ul style="list-style-type: none"> • Efficiency measures must be linked to evidence supporting one of the five clinical quality domains. • In the context of NQMC, efficiency measures typically assess the relationship of the cost of care associated with a specified level of quality of care. • These measures may address the frequency with which a less resource-intensive intervention is substituted for a more resource-intensive intervention of equal or lesser effectiveness, or a more effective intervention is substituted for a less effective intervention that is equally or more resource-intensive. • Measures in this domain may also assess the performance of activities by a health care organization or clinician to minimize waste. <p>Example: Percentage of patients with chronic periodontitis treated with scaling and root planing.</p>

Appendix 3: Sample Specification for Administrative Measures

TITLE

Description:

Numerator:

Denominator:

Rate:

Rationale:

NQF Domain:

AHRQ Domain:

IOM Aim:

Level of Aggregation:

Improvement Noted As:

Data Required:

Measure Purpose:

Applicable Stratification Variables:

Measure Limitations:

Calculation Algorithm:

Appendix 4: Request for Proposals Terms

The following should be listed as terms to the RFP.

- Neither this RFP nor any responses hereto shall be considered a binding offer or agreement. If the DQA (through the ADA) and any responding Respondent decide to pursue a business relationship for any or all of the services or equipment specified in this RFP, the parties will negotiate the terms and conditions of a definitive, binding written agreement which shall be executed by the parties. Until and unless a definitive written agreement is executed, DQA shall have no obligation with respect to any Respondent in connection with this RFP.
- This RFP is not an offer to contract, but rather an invitation to a Respondent to submit a bid. Submission of a proposal or bid in response to this RFP does not obligate the DQA to award a contract to a Respondent or to any Respondent, even if all requirements stated in this RFP are met. The DQA (through the ADA) reserves the right to contract with a Respondent for reasons other than lowest price. Any final agreement between ADA (on behalf of the DQA) and Respondent will contain additional terms and conditions regarding the provision of services or equipment described in this RFP. Any final agreement shall be a written instrument executed by duly authorized representatives of the parties.
- Respondent's RFP response shall be an offer by Respondent which may be accepted by the DQA. The pricing, terms, and conditions stated in Respondent's response must remain valid for a period of one hundred twenty (120) days after submission of the RFP to the DQA.
- This RFP and Respondent's response shall be deemed confidential DQA information. Any discussions that the Respondent may wish to initiate regarding this RFP should be undertaken only between the Respondent and DQA. Respondents are not to share any information gathered either in conversation or in proposals with any third parties, including but not limited to other business organizations, subsidiaries, partners or competitive companies without prior written permission from the DQA.
- The DQA reserves the right to accept or reject a Respondent's bid or proposal to this RFP for any reason and to enter into discussions and/or negotiations with one or more qualified Respondents at the same time, if such action is in the best interest of the DQA.

- The DQA reserves the right to select a limited number of Respondents to make a “Best and Final Offer” for the services or equipment which are the subject of this RFP. Respondents selected to provide a “Best and Final Offer” shall be based on Respondent qualifications, the submitted proposal and responsiveness as determined solely by the DQA.
- All Respondent’s costs and expenses incurred in the preparation and delivery of any bids or proposals (response) in response to this RFP are Respondent’s sole responsibility.
- Applicants should limit the Facilities & Administrative (F&A) rate not exceed 10% of the direct cost of the project.
- The DQA reserves the right to award contracts to more than one Respondent for each of the services identified in this RFP.
- All submissions by Respondents shall become the sole and exclusive property of the DQA (through the ADA) and will not be returned by the DQA or ADA to Respondents.

Appendix 5: Sample Measure Testing RFP Evaluation Summary Sheet (Administrative Data)

Proposal Number/ PI Name:

Reviewer Name:

I. **SCIENTIFIC REVIEW**

	Considerations	Strengths	Weaknesses
Data Sources	<ul style="list-style-type: none"> • Availability of Medicaid/CHIP data • Availability of commercial data • Access to patient charts for record validation • Number of states/payers represented • Diversity in provider payment mechanisms • Systems capability • Validity of coded data (Assurance of data quality) • Multi-year data set • Recent data 		
Scientific methodology	<ul style="list-style-type: none"> • Ability to assess reliability, feasibility and validity as defined in the RFP • Valid sampling methodologies if used • Descriptive statistics for the measure entities • Valid statistical tests 		
Relevant Experience	<ul style="list-style-type: none"> • Data analysis background • Record of fulfilling deliverable-based projects • Record of publications 		
Investigators	<ul style="list-style-type: none"> • Range of experience in the testing team 		
Timeline			

II. **BUDGET REVIEW** (Please enter your comments on whether the proposed expenses are justified by the methodology/ data sources proposed)

III. **SUMMARY STATEMENT:** (Impression of proposal which will be shared with the investigators as written comments from reviewers. Please keep short)

Appendix 6: NQF Criteria for Endorsement Guidance^{7,10}

Generic Rating Scale Used

Rating	Definition
High	Based on the information submitted, there is high confidence (or certainty) that the criterion is met.
Moderate	Based on the information submitted, there is moderate confidence (or certainty) that the criterion is met.
Low	Based on the information submitted, there is low confidence (or certainty) that the criterion is met.
Insufficient	There is insufficient information submitted to evaluate whether the criterion is met (e.g., blank, incomplete, or not relevant, responsive, or specific to the particular question).

1. Evidence and Performance Gap, Importance to Measure and Report Extent to which the specific measure focus is evidence-based and important to making significant gains in healthcare quality where there is variation in or overall less-than-optimal performance.

1a. Evidence to Support the Measure Focus: The measure focus is evidence-based, demonstrated as follows:

- Health outcome: a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence that the measured structure leads to a desired health outcome.
- Efficiency: evidence is required for the quality component but not required for the resource use component. (Measures of efficiency combine the concepts of resource use and quality.
- Patient-reported outcome-based performance measures (PRO-PMs): in addition to evidence required for any outcome measure, evidence should demonstrate that the target population values the measured PRO and finds it meaningful.
- Measures incorporating Appropriate Use Criteria: NQF's guidance for evidence for measures in general, and specifically those based on clinical practice guidelines, apply to measures based on appropriateness criteria as well.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

When assessing measure performance data for Performance Gap (1b), the following factors should be considered:

- distribution of performance scores;
- number and representativeness of the entities included in the measure performance data;
- data on disparities; and

¹⁰ This is a snapshot of the endorsement guidance. Please access the NQF website for a more detailed explanation of the NQF's Measure Evaluation Criteria for Endorsement.

- size of the population at risk, effectiveness of an intervention, likely occurrence of an outcome, and consequences of the quality problem.

For maintenance of endorsement: If a measure is found to be “topped out” (i.e., does not meet criteria for opportunity for improvement (1b)), the measure will be considered for inactive endorsement with reserve status only. The measure must meet all other criteria, otherwise the measure should not be endorsed.

1c. For composite performance measures, the following must be explicitly articulated and logical:

- 1c1.** The quality construct, including the overall area of quality; included component measures; and the relationship of the component measures to the overall composite and to each other; and
- 1c2.** The rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually; and
- 1c3.** How the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.

Definition /Rating	Quantity of Body of Evidence (Total number of studies (not articles or papers))	Quality of Body of Evidence (Certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence related to study factors including: study design or flaws; directness/indirectness to the specific measure (regarding the population, intervention, comparators, outcomes); imprecision (wide confidence intervals due to few patients or events))	Consistency of Results of Body of Evidence (Stability in both the direction and magnitude of clinically/practically meaningful benefits and harms to patients (benefit over harms) across studies in the body of evidence)
High	5+ studies	Randomized controlled trials (RCTs) providing direct evidence for the specific measure focus, with adequate size to obtain precise estimates of effect, and without serious flaws that introduce bias	Estimates of clinically/practically meaningful benefits and harms to patients are consistent in direction and similar in magnitude across the preponderance of studies in the body of evidence
Moderate	2-4 studies	<ul style="list-style-type: none"> • Non-RCTs with control for confounders that could account for other plausible explanations, with large, precise estimate of effect <p>OR</p> <ul style="list-style-type: none"> • RCTs without serious flaws that introduce bias, but with either indirect evidence or imprecise estimate of effect 	Estimates of clinically/practically meaningful benefits and harms to patients are consistent in direction across the preponderance of studies in the body of evidence, but may differ in magnitude. If only 1 study, then the estimate of benefits greatly outweighs the estimate of potential harms to patients (1 study cannot achieve high consistency rating)

Low	1 study	<ul style="list-style-type: none"> • RCTs with flaws that introduce bias OR <ul style="list-style-type: none"> • Non-RCTs with small or imprecise estimate of effect, or without control for confounders that could account for other plausible explanations 	<ul style="list-style-type: none"> • Estimates of clinically/practically meaningful benefits and harms to patients differ in both direction and magnitude across the preponderance of studies in the body of evidence OR <ul style="list-style-type: none"> • wide confidence intervals prevent estimating net benefit <p>If only 1 study, then estimated benefits do not greatly outweigh harms to patients</p>
Insufficient to Evaluate	<ul style="list-style-type: none"> • No empirical evidence OR <ul style="list-style-type: none"> • Only selected studies from a larger body of evidence 	<ul style="list-style-type: none"> • No empirical evidence OR <ul style="list-style-type: none"> • Only selected studies from a larger body of evidence 	No assessment of magnitude and direction of benefits and harms to patients

2. Reliability and Validity—Scientific Acceptability of Measure Properties: Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.

2a. Reliability Use

- **2a1.** The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. For any measures that use ICD-9 CM codes, ICD-10 CM codes must also be provided. If HHS implements ICD-10 as planned in October 2015, then NQF will no longer accept ICD-9 CM codes for measures after December 31, 2015. eMeasures should be specified in the Health Quality Measures Format (HQMF) and must use the Quality Data Model (QDM) and value sets vetted through the National Library of Medicine's Value Set Authority Center (VSAC). Specifications for PRO-PMs also include specific PROM(s); standard methods, modes, and languages of administration; whether (and how) proxy responses are allowed; standard sampling procedures; handling of missing data; and calculation of response rates to be reported with the performance measure results. Specifications for composite performance measures include component measure specifications (unless individually endorsed); aggregation and weighting rules; handling of missing data; standardizing scales across component measures; required sample sizes.
- **2a2.** Reliability testing demonstrates that the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For PRO-PMs and composite performance measures, reliability should be demonstrated for the computed performance score.

2b. Validity Use

- **2b1.** The measure specifications are consistent with the evidence presented to support the focus of measurement under criterion 1a. The measure is specified to capture the most inclusive target population indicated by the evidence, and exclusions are supported by the evidence.
 - **2b2.** Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.
 - **2b3.** Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
- AND**

- If patient preference (e.g., informed decision making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).
- **2b4.** For outcome measures and other measures when indicated (e.g., resource use):
an evidence-based risk-adjustment strategy is specified; is based on patient factors (including clinical and sociodemographic risk factors) that influence the measured outcome and are present at start of care; and has demonstrated adequate discrimination and calibration.

OR

rationale/data support no risk adjustment.

- **2b5.** Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance;

OR

there is evidence of overall less-than-optimal performance.

- **2b6.** If multiple data sources/methods are specified, there is demonstration they produce comparable results.
- **2b7.** For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

2c. Disparities (*Disparities should be addressed under subcriterion 1b*)

If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);

2d. For composite performance measures, empirical analyses support the composite construction approach and demonstrate the following: H M L I

- **2d1.** the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and
- **2d2.** the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

3. Feasibility: Extent to which the specifications, including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- **3a.** For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).
- **3b.** The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.
- **3c.** Demonstration that the data collection strategy (e.g., data source/availability, timing, frequency, sampling, patient-reported data, patient confidentiality costs associated with fees/licensing for proprietary measures or elements such as risk model, grouper, instrument) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).
- For eMeasures, a feasibility assessment is required; this feasibility assessment must address the data elements and measure logic and demonstrate that the eMeasure can be implemented or that feasibility concerns can be adequately addressed. The feasibility assessment uses a standard score card or a fully transparent alternative that includes at a minimum: 1) a description of the assessment, feasibility scores for all data elements, and explanatory notes for all

data element components scoring a "1" (lowest rating); 2) demonstration that the measure logic can be executed; w and 3) plan for addressing feasibility concerns.

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

- **4a. Accountability and Transparency**

Performance results are used in at least 1 accountability application within 3 years after initial endorsement and are publicly reported within 6 years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

AND

- **4b. Improvement**

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

AND

4c. The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence