DQA

DENTAL QUALITY ALLIANCE™

Improving Oral Health Through Measurement

PROCEDURE MANUAL FOR PERFORMANCE MEASURES DEVELOPMENT AND MAINTENANCE

A Voluntary Consensus Process

NOVEMBER 2022

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

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

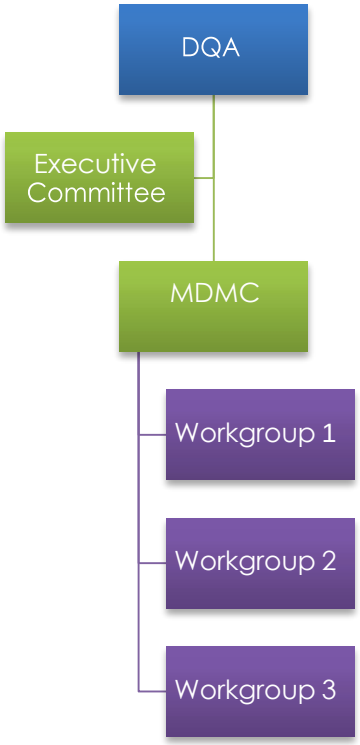
1. To identify and develop evidence-based oral health care performance measures and measurement resources.
2. To advance the effectiveness and scientific basis of clinical performance measurement and improvement.
3. To foster and support professional accountability, transparency, and value in oral health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

Measure Development and Maintenance Committee (MDMC)

The Measure 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, then develop and test 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 Guides. 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 nominates subject matter experts to the MDMC and the workgroups. Subject matter experts should be (1) capable of knowledgeably participating in the measure development activities; (2) able to work collaboratively in a small group; 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. 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

Quality Concepts

Quality is the "degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge."[1]

The ultimate desired health outcome is optimum oral health representing the "ability to speak, smile, smell, taste, touch, chew, swallow and convey a range of emotions through facial expressions with confidence and without pain, discomfort and disease of the craniofacial complex."[2][3] The core elements of oral health are as follows: *disease and condition status* "refers to a threshold of severity or a level of progression of disease, which also includes pain and discomfort"; *physiological function* "refers to the capacity to perform a set of actions that include, but are not limited to, the ability to speak, smile, chew, and swallow"; and *psychosocial function* "refers to the relationship between oral health and mental state that includes, but is not limited to, the capacity to speak, smile, and interact in social and work situations without feeling uncomfortable or embarrassed."[3]

The goal of improving quality is to deliver safe, effective, efficient, patient-centered, timely and equitable care.[4]

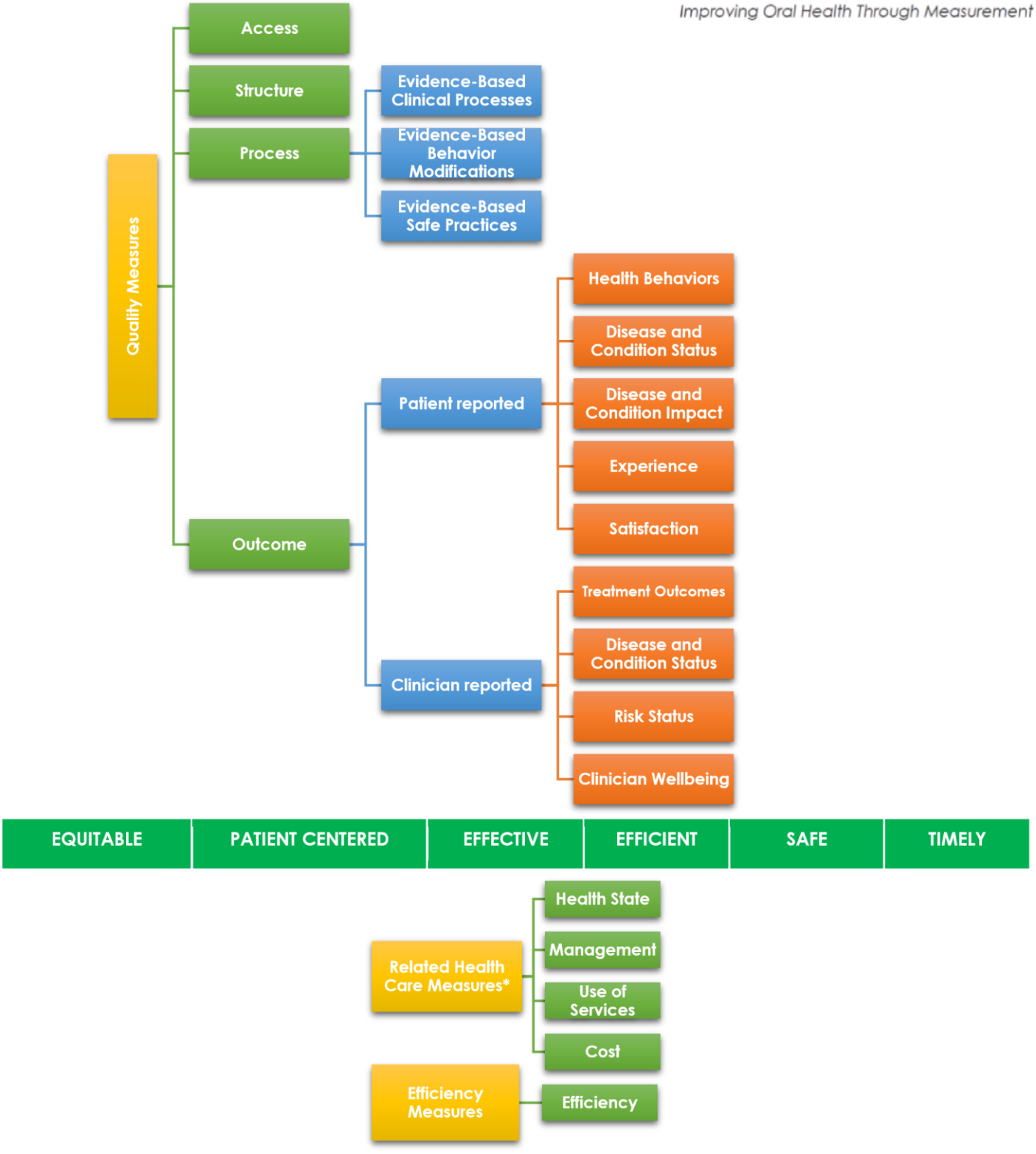
[1] Institute of Medicine (US) Committee to Design a Strategy for Quality Review and Assurance in Medicare; Lohr KN, editor. Washington (DC): National Academies Press (US); 1990. Accessed at https://www.ncbi.nlm.nih.gov/books/NBK235472/ on September 23, 2021.

[2] Glick M, Williams DM, Kleinman DV, Vujicic M, Watt RG, Weyant RJ. A new definition for oral health developed by the FDI World Dental Federation opens the door to a universal definition of oral health. J Public Health Dent. 2017 Dec; 77(1):3-5

[4] Institute of Medicine. 2001. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: The National Academies Press.

DQA Measure Development Framework to Measure Health and Healthcare Quality –

DQA
DENTAL QUALITY ALLIANCE™
Improving Oral Health Through Measurement



The DQA has historically looked at the work of the Institute of Medicine (IOM), National Quality Measurement Clearnghouse (NQMC) and National Quality Forum (NQF) as the athoritative voices in the quality measurement landscape.

In November 2021, the DQA expanded on the NQMC domains with this updated framework[4] to identify measurement gaps and support measure prioritization activities. The intent of this framework is to anchor current DQA measures and serve as the guide for the identification of future measure development activities. It is not intended to replace industry-accepted frameworks. **Appendix 2** presents the detailed definitions for each of the terms included in this framework.

[4] Adapted from National Quality Measures Clearinghouse. NQMC Measure Domain Framework. July 2018; https://www.ahrq.gov/gam/summaries/domain-framework/index.html and
Cochrane Effective Practice and Organisation of Care (EPOC). *What Outcomes Should be Reported in EPOC Reviews.* EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed September 23, 2021)

Measure Development Process

The process of developing measures (Figure 1) 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

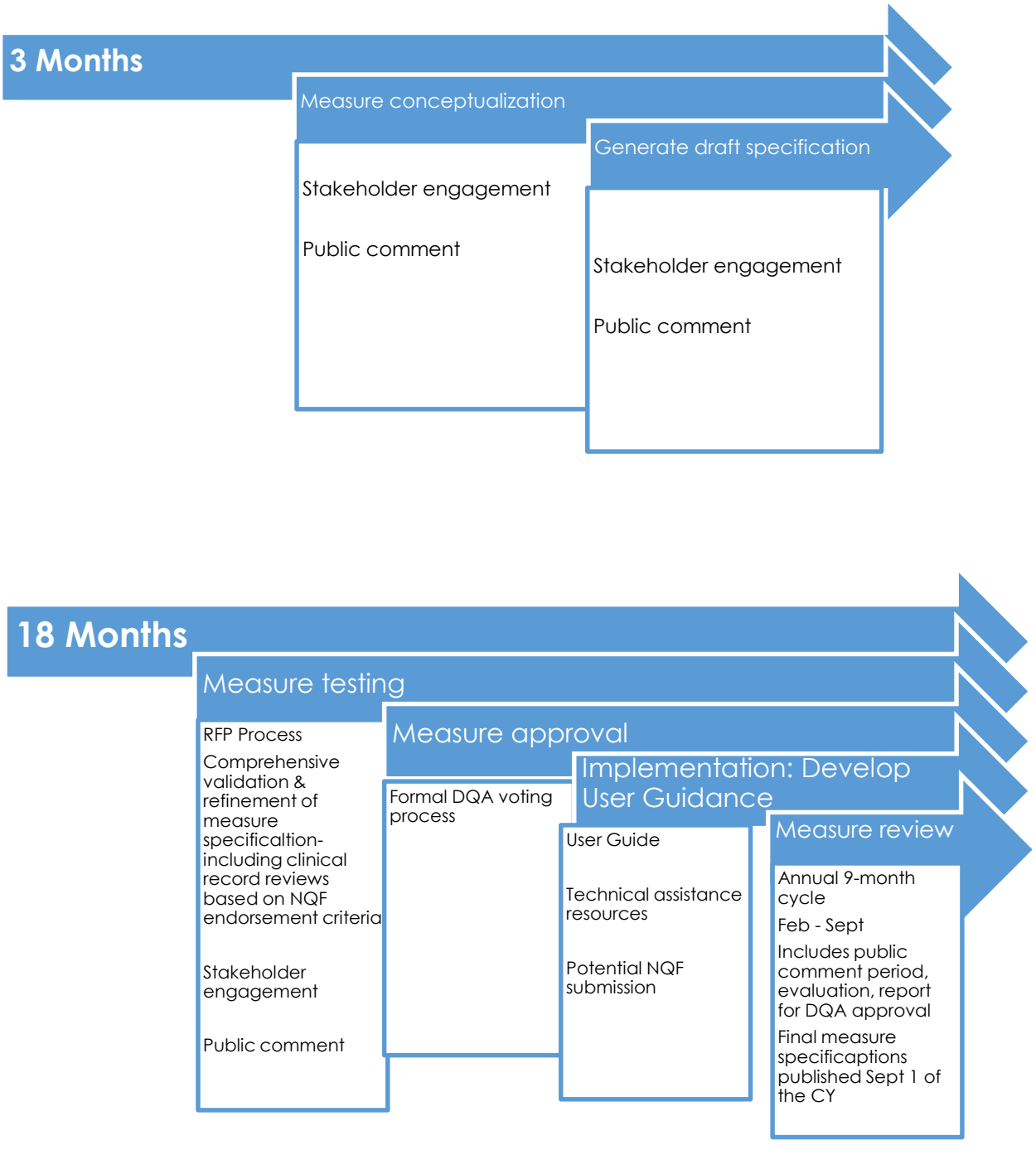
- Developing a workplan for measure testing
- Overseeing and guiding feasibility, reliability and validity testing
- Issuing an Interim Report of testing results and a public call for comment
- DQA review and approval of fully specified and finalized measures
- Developing the Final Report

Phase 3: Measure Dissemination

- Issue Final Report, with revised specifications and User Guides
- 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.

Figure 1: DQA Measure Development – Process Overview/Projected Timeline



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. Several comprehensive scans are published on the DQA website, including: a 2020 scan of patient reported oral health measures, a 2017 scan on oral health quality improvement initiatives, a 2015 scan on practice-based measures, and the original 2012 scan of pediatric measures. The MDMC will continue to undertake environmental scans, as appropriate and when needed, in support of its work.

Environmental scan resources include:

1. PubMed searches – The National Library of Medicine has deployed a search interface to find citations relating to healthcare quality
2. Keyword searches of the internet using standard search engines such as google
3. Soliciting measures from other measure development organizations (e.g., Veterans Administration, public and private payers, HRSA programs)
4. 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.[5] Criteria for this rating exercise for measure concepts shall be based on those established by the National Advisory Council for Healthcare Research and Quality Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP, estabilished by the Agency for Healthcare Research and Quality (AHRQ).[6]

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 NQF process identifies the following considerations when evaluating measure concept **importance**, **validity**, and **feasibility**.[7]

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;
3. Health care systems should clearly be accountable for the quality problem assessed

[5] 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

[6] 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.

[7] 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.

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).

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) without undue burden in implementation.
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 framework developed by the DQA that is adapted from the NQMC Measure Domain Framework (NQMC) (**Appendix 2**).

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) 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

MDMC conducts extensive testing to establish feasibility, validity, reliability, and usability. Availability of data source needed for measures evaluation is the first step in determining if the testing will be conducted internally or DQA may issue a request for proposals (RFP) to identify a testing partner.

The DQA may use a competitive 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;[8]
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

[8] 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.

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 NQF. The NQF requires data for topic importance, performance gap, evidence to support process measures, scientific soundness, feasibility, and use and 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.[9] 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 following[10]:

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. Specific questions to guide new measure development:

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?

[9] Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement. September 2021. Accessed at: https://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx

[10] Adapted from the National Quality Measures Clearinghouse Tutorials on Quality Measures: http://www.qualitymeasures.ahrq.gov/tutorial/index.aspx ; resource sunset in 2018 National Quality Forum. Measure Evaluation Criteria:https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88439. Accessed September 2022

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 - 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. Specific questions to guide new measure development:

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.

Validity - 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. Specific questions to guide new measure development:

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 - Extent to which potential audiences (e.g., consumers, purchasers, providers, and 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. Specifically:

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 and a final vote for approval.

Developing the final report

After the DQA has voted to approve a measure, a final report is prepared.

It 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 are documented. The report addresses 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 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.

Phase 1: Call for Comments

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

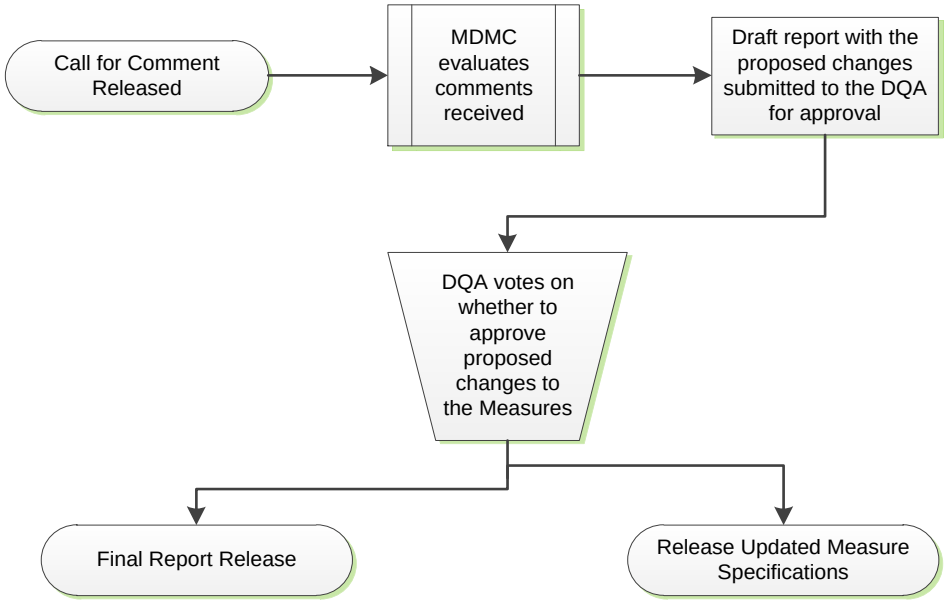
Phase 2: Review & Evaluation

- 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 Guides

Phase 3: Approval

- Voting by the DQA on proposed changes

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 1st - April 30th, 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 Guides voted on by the DQA
September 1, 20XX	• Final Report of the Annual Measure Review Released • Updated measure specifications released/disseminated via DQA website • Updated User Guides released/disseminated via DQA website
January 1st, 20XX	The updated versions effective January 1st 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 American Dental Association holds copyright on behalf of the Dental Quality Alliance. 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 American Dental Association (ADA) may include creating, or contributing to the creation of, original content for one or more of the ADA's ongoing publications or for a special project that may result in a publication distributed by the ADA.

The undersigned irrevocably grants, assigns, and transfers to the ADA 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 American Dental Association as "works made for hire". The undersigned understand that he or she is acting as an independent contractor respecting volunteer work performed for the ADA, 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 ADA'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: Defining Terms From Across the Quality Landscape

Healthcare Delivery Measures

These are 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.

Population Health Measures

These are applied to groups of persons identified by geographic location, organizational affiliation or non-clinical characteristics, in order to assess public health programs, community influences on health, or population-level health characteristics that may not be directly attributable to the care delivery system.

Defining Terms From Across the Quality Landscape:	
Aims for Improvement[11]	**Equitable:** providing care that does not vary in quality because of personal characteristics, such as gender, age, race, ethnicity, education, disability, sexual orientation, geographic location, and socioeconomic status
	Patient-centered: providing care that is respectful of, and responsive to, individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions
	Effective: providing services based on scientific knowledge to all who could benefit and refraining from providing services to those not likely to benefit (avoiding underuse and overuse)
	Efficient: avoiding waste, in particular waste of equipment, supplies, ideas, and energy
	Safe: avoiding injuries to patients from the care that is intended to help them
	Timely: reducing waits and sometimes harmful delays for both those who receive and those who give care
Quality Measures[12]: 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.	

[11] Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US); 2001. PMID: 25057539.
[12] NQMC Measure Domain Definitions. Content last reviewed July 2018. Agency for Healthcare Research and Quality, Rockville, MD. https://www.ahrq.gov/gam/summaries/domain-definitions/index.html. Accessed September 2022

<table>
<tr><td colspan="2">Access: Access to care is the attainment of timely and appropriate health care by patients or enrollees of a health care organization or clinician (or of a public health intervention by a population). Access measures are supported by evidence that an association exists between the measure and the outcomes of or satisfaction with care.</td></tr>
<tr><td colspan="2">Structure: Structure of care is a feature of a health care organization or clinician (or public health program for populations) related to the capacity to provide high quality health care. Structure measures are supported by evidence that an association exists between the measure and one of the other clinical quality measure domains.</td></tr>
<tr><td rowspan="3">Process; A process of care is a health care-related activity performed for, on behalf of, or by a patient. Process measures are supported by evidence that the process—that is the focus of the measure—has led to improved outcomes.</td><td>Evidence-Based Clinical Processes: Oral health care is provided using the judicious integration of systematic assessments of clinically relevant scientific evidence (evidence-based guidelines), relating to the person's oral and medical condition and history, with the oral health care provider's clinical expertise and the person's treatment needs and preferences.[13]</td></tr>
<tr><td>Evidence-Based Behavior Modifications: Evidence-Based interventions to positively influence health behaviors. Behavior change requires attention to individuals (e.g., personal health behaviors), families (e.g., family stress, social support), health care professionals (e.g., appropriate counseling techniques), the environment (e.g., accessibility to oral health care, status of community water fluoridation), and cross-cutting issues (e.g., racial and ethnic health disparities, cultural preferences).[14]</td></tr>
<tr><td>Evidence-Based Safe Practices: The evidence-based safe practices are ready-to-use tools to improve safety and have been evaluated, assessed and endorsed to guide large and small healthcare systems in providing the safest care possible.[15]</td></tr>
<tr><td colspan="2">Outcome: An outcome of care is a health state of a patient resulting from health care. Outcome measures are supported by evidence that the measure has been used to detect the impact of one or more clinical interventions (or public health interventions for population outcomes). Measures in this domain are attributable to antecedent health care (or public health interventions) and should include provisions for risk-adjustment. Outcomes may be reported by patients or clinician assessed.</td></tr>
</table>

[13] ADA Policy Statement on Evidence-based Dentistry. http://www.ada.org/en/about-the-ada/ada-positions-policies-and-statements/policy-on-evidence-based-dentistry. Accessed May 13th, 2018

[14] IOM (Institute of Medicine). 2011. *Advancing Oral Health in America*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13086.

[15] National Quality Forum (NQF). Safe Practices for Better Healthcare. https://www.qualityforum.org/News_And_Resources/Press_Kits/Safe_Practices_for_Better_Healthcare.aspx Accessed September 23, 2021

Patient Reported Outcomes: "Reports of the patient's health status, health behavior, experience with health care or satisfaction with health care that comes directly from the patient"[16] as a result of healthcare structures and processes and are supported by evidence that the healthcare system can influence the outcome.
Health Behaviors: These are "actions taken by individuals that affect health or mortality, may be intentional or unintentional, and can promote or detract from the health of the actor or others. Examples include smoking, substance use, diet, physical activity, sleep, risky sexual activities, health care seeking behaviors, and adherence to prescribed medical treatments."[17]
Oral Health Status: "The ability to speak, smile, smell, taste, touch, chew, swallow and convey a range of emotions through facial expressions with confidence and without pain, discomfort and disease of the craniofacial complex."**Error! Bookmark not defined.** **Disease and condition status:** Measures of disease and condition status address diseases of the craniofacial structures: e.g., caries status, tooth loss, and bleeding gums. Measures of condition and disease status often are more reliably assessed through clinical evaluations. When advancing severity of disease is inferred from procedure codes, appropriate validation testing must be conducted to determine if the measure can be classified in this domain. Any concerns related to confounding by access or difficulties in accurately identifying disease severity on the basis of procedure codes alone without diagnoses codes should be evaluated.**Error! Bookmark not defined.** But patient-reported indicators may be important when clinical assessments are not available or as a gauge of a patient's understanding and perception of his/her oral health status. **Disease and condition impact**: Refers to "patient-perceived impact of oral conditions and dental interventions"[18] and include pain, appearance (aesthetics), functional status and psychosocial impacts.
Patient Experience: Experience of care is a patient's or enrollee's report of observations of and participation in health care, or assessment of any resulting change in their health. Patient experience measures are supported by evidence that an association exists between the measure and patients' values and preferences, or one of the other clinical quality domains.
Patient Satisfaction: "Satisfaction is about whether a patient's expectations about a health encounter were met. Two people who receive the exact same care, but who have different expectations for how that care is supposed to be delivered, can give different satisfaction ratings because of their different expectations."
Treatment Outcomes: Anticipated and unanticipated complications and consequences, as well as functional, physiological and aesthetic outcomes of care.

[16] National Quality Forum. Patient Reported Outcomes in Performance Measurement. 2013; http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72549. Accessed September 23, 2021

[17] Short, S. E., & Mollborn, S. (2015). Social Determinants and Health Behaviors: Conceptual Frames and Empirical Advances. *Current opinion in psychology*, *5*, 78–84. doi:10.1016/j.copsyc.2015.05.002

[18] John, M. T., Feuerstahler, L., Waller, N., Baba, K., Larsson, P., Celebić, A., Kende, D., Rener-Sitar, K., & Reissmann, D. R. (2014). Confirmatory factor analysis of the Oral Health Impact Profile. *Journal of oral rehabilitation*, *41*(9), 644–652. https://doi.org/10.1111/joor.12191

<table>
<tr><td colspan="2">Risk Status: There are patient-related attributes or characteristics that contribute to outcomes[19]. (Examples include patient's primary diagnosis and condition severity, comorbid conditions, genetic, biological, demographic, socioeconomic, environmental, and psychosocial factors; health-related behaviors; and attitudes, preferences and perceptions regarding health care). These are collectively termed risk factors and understanding any changes in the risk status influences patient's oral health and potential outcomes of care.</td></tr>
<tr><td colspan="2">Clinician Wellbeing: Clinician well-being is essential for safe, high-quality patient care and supports improved patient-clinician relationships, a high-functioning care team, and an engaged and effective workforce.[20]</td></tr>
<tr><td rowspan="4">Related Healthcare 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.</td><td>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.</td></tr>
<tr><td>Management: 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. Management measures assess administrative activities that are important to health care but are not part of the direct interaction between individual patients and health care professionals.</td></tr>
<tr><td>Use of Services: 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.</td></tr>
<tr><td>Cost: 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.</td></tr>
<tr><td>Efficiency Measures</td><td>Efficiency: 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 population health processes are implemented that have been demonstrated by evidence to be efficient).</td></tr>
</table>

[19] National Quality Forum. Glossary of Terms. Accessed on: http://www.qualityforum.org/Measuring_Performance/Measuring_Performance.aspx

[20] National Academy of Medicine. Action Collaborative on Clinician Well-being and Resilience. Accessed from: https://nam.edu/initiatives/clinician-resilience-and-well-being/ Accessed September 23, 2021

Appendix 3: Sample Specification for Administrative Measures

TITLE

Description:
Numerator:
Denominator:
Rate:

Rationale:
NQF 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 to 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	• 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	• 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	• Data analysis background • Record of fulfilling deliverable-based projects • Record of publications		
Investigators	• 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[21]

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:

- **Outcome**: Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **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.)
- For measures derived from **patient reports**, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

Process measures incorporating Appropriate Use criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply 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, such as by age, sex, race, ethnicity, geography, disability, and insurance status.

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;

[21] 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; September 2021. https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88439. Accessed September 2022.

- data on disparities; and
- 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.
- Under NQF's revised approach to the evaluation of currently endorsed measures, there is a shift in emphasis for several of the evaluation criteria/subcriteria. For performance gap, there is increased emphasis on current performance and opportunity for improvement. Measure stewards are expected to provide current performance data. If limited data are available (e.g., use is voluntary), data from the literature can be considered.

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	• Non-RCTs with control for confounders that could account for other plausible explanations, with large, precise estimate of effect **OR** • 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	• RCTs with flaws that introduce bias **OR** • Non-RCTs with small or imprecise estimate of effect, or without control for confounders that could account for other plausible explanations	• 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** • wide confidence intervals prevent estimating net benefit If only 1 study, then estimated benefits do not greatly outweigh harms to patients
Insufficient to Evaluate	• No empirical evidence **OR** • Only selected studies from a larger body of evidence	• No empirical evidence **OR** • 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. –

- Measure specifications include the target population (denominator) to whom the measure applies, identification of those from the target population who achieved the specific measure focus (numerator, target condition, event, outcome), measurement time window, exclusions, risk adjustment/stratification, definitions, data source, code lists with descriptors, sampling, scoring/computation.
- All measures that use the ICD classification system must use ICD-10-CM.
- eCQMs should be specified using the latest industry accepted eCQM technical specifications: health quality measure format (HQMF), Quality Data Model (QDM), Clinical Quality Language (CQL), and value sets vetted through the National Library of Medicine's Value Set Authority Center (VSAC).
- Specifications for instrument-based measures also include the specific instrument (e.g., 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.
- Under NQF's revised approach to the evaluation of currently endorsed measures, there is a shift in emphasis for several of the evaluation criteria/subcriteria. However, there is no change in the evaluation of the current specifications.

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 **instrument-based measures** (including PRO-PMs), reliability must be demonstrated for the data element level as well as for the computed performance score. For **composite performance measures**, reliability must be demonstrated for the computed performance score.

- Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to, inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
- Testing must be conducted for the measure as specified (e.g., all relevant levels of analysis, using applicable data sources, care settings, patients, providers, etc.). If more than one measure is included under one NQF number, each measure must be tested per NQF evaluation requirements. If more than one level of analysis is specified, testing must be conducted for each level separately.
- Testing at the level of data elements requires that all critical data elements be tested (not just agreement of one final overall computation for all patients). At a minimum, the numerator, denominator, and exclusions (or exceptions) must be assessed and reported separately.
- For accountable-entity level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred.

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 **instrument-based measures** (including **PRO-PMs**), validity must be demonstrated for the data element level as well as for the computed performance score. For **composite performance measures**, validity must be demonstrated for the computed performance score by the time of endorsement maintenance; if empirical testing of the computed performance score is not feasible at the time of initial endorsement, acceptable alternatives include systematic assessment of content or face validity of the composite performance measure or demonstration that each of the component measures meet NQF subcriteria for validity (via either empirical testing of the data elements or measure score or via face validity).
- Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to, testing hypotheses that the measures scores indicate quality of care (e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures).
- Testing must be conducted for the measure as specified (e.g., all relevant levels of analysis, using applicable data sources, care settings, patients, providers, etc.). If more than one measure is included under one NQF number, each measure must be tested per NQF evaluation requirements. If more than one level of analysis is specified, testing must be conducted for each level separately
- Testing at the level of data elements requires that all critical data elements be tested (not just agreement of one final overall computation for all patients). At a minimum, the numerator, denominator, and exclusions (or exceptions) must be assessed and reported separately
- If presenting score-level validation (typically via construct validity or known-groups analysis) the following should be included
-
- ○ Narrative describing the hypothesized relationships
- ○ Narrative describing why examining these relationships (e.g., correlating measures) would validate the measure
- ○ Expected direction of the association
- ○ Expected strength of the association
- ○ Specific statistical tests used (more detail is better)
- ○ Results of the analysis
- ○ Interpretation of those results (including how they related to the hypothesis and whether they have helped to validate the measure)
- **2b3.** Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure.

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, cost):
  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.** 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. - Examples of evidence that missing data distorts measure results include, but are not limited to, frequency of occurrence and variability across measured entities.

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 (EHRs) or other electronic sources. If the required data are not in EHRs 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 eMeasure feasibility assessment report and scorecard were updated in 2013 and can be accessed here: http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=73039

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.

- **4a1. Accountability and Transparency**
  Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within sixyears 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.
  4a2. Feedback on the measure by those being measured or others is demonstrated when: 1. those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data 2. those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation 3. this feedback has been considered when changes are incorporated into the measure
  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
  4b2. The **benefits** of the performance measure in facilitating progress toward achieving high quality, efficient healthcare for individuals or populations **outweigh evidence of unintended negative consequences** to individuals or populations (if such evidence exists).
  4b3. Data and result detail are maintained such that the resource use measure, including the clinical and construction logic for a defined unit of measurement, can be deconstructed to facilitate transparency and understanding.