NATIONAL QUALITY FORUM

Primary Care and Chronic Illness, Fall 2018 Measure Review Cycle

Standing Committee Measure Evaluation Web Meetings

Samuel Stolpe, PharmD, MPH
Suzanne Theberge, MPH
Hiral Dudhwala, RN, MSN/MPH

February 4, 2019

Welcome

Project Team – Primary Care and Chronic Illness

Samuel Stolpe, PharmD, MPH
Senior Director

Suzanne Theberge, MPH
Senior Project Manager

Hiral Dudhwala, RN, MSN, MPH
Project Manager

Agenda

- Welcome
- Introductions and Disclosure of Interest
- Overview of Evaluation Process
- Consideration of Two Candidate Measures
- Public Comment
- Consideration of Related and Competing Measures
- Next Steps

Introductions and Disclosures of Interest

Primary Care and Chronic Illness Fall 2018 Cycle Standing Committee

*New Committee Members

- **Dale Bratzler, DO, MPH (Co-Chair)**
- **Adam Thompson, BA (Co-Chair)**
- Thiru Annaswamy, MD
- Robert Bailey, MD
- Lindsay Botsford, MD
- Roger Chou, MD
- William Curry, MD, MS
- Jim Daniels, BSN
- Woody Eisenberg, MD*
- Kim Elliott, PhD
- V. Katherine Gray, PhD
- Starlin Haydon-Greatting, MS, BS, Pharm, FAPhA*
- Ann Kearns, MD, PhD
- Anne Leddy, MD, FACE
- Grace Lee, MD
- Anna McCollister-Slipp
- Janice Miller, DNP, CRNP, CDE
- James Rosenzweig, MD*
- Steven Strode, MD, Med, MPH, FAAFP
- William Taylor, MD
- Kimberly Templeton, MD
- John Ventura, DC

Primary Care and Chronic Illness
Expert Reviewers

- Amesh Adalja, MD
- Esther Babady, PhD, (ABMM)
- Carlos Bagley, MD, FAANS
- Kenneth Benson
- Tamala Bradham, DHA, PhD, CCC-A
- Kathleen Brady, MD, MSCE
- Steven Brotman, MD, JD
- Craig Butler, MD, MBA, CPE
- Laura Evans, MD, MSc
- Scott Friedman, MD
- Piero Garzaro, MD
- William Glomb, MD, FCCP, FAAP
- Donald Goldmann, MD
- Stephen Grossbart, PhD
- James Mitchell Harris, PhD
- Marci Harris Hayes, PT, DPT, MSCI, OCS
- Jeffrey Hart, MS
- Mark Jarrett, MD, MBA

Primary Care and Chronic Illness
Expert Reviewers

- Michael Lane, MD, MSc, MPHS, CPPS
- David Lang, MD
- Jeffrey Lewis, BA
- Jason Matuszak, MD, FAAFP, CAQSM, RMSK
- John McClay, MD
- Richard Murray, MD
- Melinda Neuhauser, PharmD, MPH, FCCP, FASHP
- Rocco Orlando, MD, FACS
- Crystal Riley, PharmD, MHA, MBA, CPHQ, CHPIT
- Catherine Roberts, MD
- Christine Schindler, PhD, RN, CPNP-AC/PC, WCC
- Christopher Visco, MD
- Jacquelyn Youde, AuD, CCC-A

Overview of Evaluation Process

Roles of the Standing Committee *During the Evaluation Meeting*

- Act as a proxy for the NQF multistakeholder membership
- Work with NQF staff to achieve the goals of the project
- Evaluate each measure against each criterion
  - *Indicate the extent to which each criterion is met and rationale for the rating*
- Make recommendations regarding endorsement to the NQF membership
- Oversee portfolio of Primary Care and Chronic Illness measures

Ground Rules for Today's Meeting

During the discussions, Committee members should:

- Be prepared, having reviewed the measures beforehand
- Base evaluation and recommendations on the measure evaluation criteria and guidance
- Remain engaged in the discussion without distractions
- Attend the meeting at all times
- Keep comments concise and focused
- Avoid dominating a discussion and allow others to contribute
- Indicate agreement without repeating what has already been said

Process for Measure Discussion and Voting

- Brief introduction by measure developer (3-5 minutes)
- Lead discussants will begin Committee discussion for each criterion:
  - *Briefly explaining information on the criterion provided by the developer*
  - *Providing a brief summary of the pre-meeting evaluation comments*
  - *Emphasizing areas of concern or differences of opinion*
  - *Noting, if needed, the preliminary rating by NQF*
    - » This rating is intended to be used as a guide to facilitate the Committee's discussion and evaluation.
- Developers will be available to respond to questions at the discretion of the Committee
- Full Committee will discuss, then vote on the criterion, if needed, before moving on to the next criterion

Endorsement Criteria

- **Importance to measure and report**: Goal is to measure those aspects with greatest potential of driving improvements; if not important, the other criteria are less meaningful (must-pass)
- **Reliability and Validity-Scientific Acceptability of measure properties**: Goal is to make valid conclusions about quality; if not reliable and valid, there is risk of improper interpretation (must-pass)
- **Feasibility**: Goal is to, ideally, cause as little burden as possible; if not feasible, consider alternative approaches
- **Usability and Use**: Goal is to use for decisions related to accountability and improvement; if not useful, probably do not care if feasible
- **Comparison to related or competing measures**

Voting on Endorsement Criteria

- **Importance to Measure and Report (must-pass):**
  - Vote on evidence (if needed), performance gap, and for composite, quality construct and rationale
- **Scientific Acceptability (must pass)**:
  - Vote on reliability and validity (if needed), and for composite, the quality construct
- **Feasibility**:
  - Vote on feasibility
- **Usability and Use (Use is must-pass for maintenance measures)**:
  - Vote on usability and use
- **Overall Suitability for Endorsement**

If a measure fails on one of the must-pass criteria, there is no further discussion or voting on the subsequent criteria for that measure; we move to the next measure.

Achieving Consensus

- Quorum: 66% of the Committee
- Pass/Recommended: Greater than 60% “Yes” votes of the quorum (this percent is the sum of high and moderate)
- Consensus not reached (CNR): 40-60% “Yes” votes (inclusive of 40% and 60%) of the quorum
- Does not pass/Not Recommended: Less than 40% “Yes” votes of the quorum
- CNR measures move forward to public and NQF member comment and the Committee will revote

Committee Quorum and Voting

- If at any point quorum is lost during a meeting, the Committee will continue the measure discussion but will not vote during the meeting. Following the meeting, staff will only send the voting survey to those Committee members who participated in the meeting. Staff will not request votes from any Committee member who did not attend the meeting.
- If staff have to request votes from the Committee following the meeting, Committee member votes must be submitted within 48 hours of the meeting.
- If a Committee member leaves the meeting and quorum is still present, the Committee can continue to vote on the measures. The Committee member who left the meeting does not need to vote on the missed measures.

Fall 2018 Cycle Measures

Two Measures for Committee Review

- 0729 Optimal Diabetes Care (Maintenance Measure)
- 3475e Appropriate Use of DXA Scans in Women Under 65 Years Who Do Not Meet the Risk Factor Profile for Osteoporotic Fracture (New Measure/eMeasure)

NQF Scientific Methods Panel Review

- The Scientific Acceptability section of this measure was independently evaluated by the Scientific Methods Panel.
  - 0729 Optimal Diabetes Care (Composite Measure)

- The Panel, consisting of individuals with methodologic expertise, was established to help ensure a higher-level evaluation of the scientific acceptability of complex measures.

Questions?

Voting Process

Consideration of Candidate Measures

0729 Optimal Diabetes Care

- **Measure Developer:** Minnesota Community Measurement
- Maintenance Measure
- Composite Measure
  - *Reviewed by NQF Scientific Methods Panel for Scientific Acceptability Criterion*

- **Measure Description:** The percentage of patients 18-75 years of age who had a diagnosis of type 1 or type 2 diabetes and whose diabetes was optimally managed during the measurement period as defined by achieving ALL of the following:
  - *HbA1c less than 8.0 mg/dL*
  - *Blood Pressure less than 140/90 mmHg*
  - *On a statin medication, unless allowed contraindications or exceptions are present*
  - *Non-tobacco user*
  - *Patient with ischemic vascular disease is on daily aspirin or anti-platelets, unless allowed contraindications or exceptions are present*

3475e Appropriate Use of DXA Scans in Women Under 65 Years Who Do Not Meet the Risk Factor Profile for Osteoporotic Fracture

- **Measure Steward:** Centers for Medicare & Medicaid Services
- **Measure Developer:** National Committee for Quality Assurance
- New Measure
- eMeasure
- **Measure Description:** Percentage of female patients 50 to 64 years of age without select risk factors for osteoporotic fracture who received an order for a dual-energy x-ray absorptiometry (DXA) scan during the measurement period.

Public Comment

Next Steps

Next Steps for Fall 2018 Cycle

- Measure Evaluation Web Meetings
  - *February 4, 2019 1-3 pm EST*
  - *February 5, 2019 2-4 pm EST*
- Post-Measure Evaluation Web Meeting
  - *February 11, 2019 2-4 pm EST (CANCELLED)*

- Draft Report Comment Period (30 days)
  - *March 18-April 16, 2019 (tentative)*
- Committee Post-Comment Web Meeting
  - *May 6, 2019 2-4 pm EST*

Spring 2019 Cycle Updates

- Intent to submit deadline – January 7, 2019
- 12 measures submitted
  - *11 maintenance measures*
  - *1 new measure*
  - *3 measures considered "complex" and will be reviewed by NQF Scientific Methods Panel for scientific acceptability criterion.*

Project Contact Info

- Email: primarycare@qualityforum.org
- NQF phone: 202-783-1300
- Project page: http://www.qualityforum.org/Primary_Care_and_Chronic_Illness.aspx
- SharePoint site: http://share.qualityforum.org/Projects/Primary%20Care%20and%20Chronic%20Illness/SitePages/Home.aspx

THANK YOU

Additional Slides, if needed

Criterion #1: Importance to Measure and Report (page 31-39)

1. 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: *the measure focus is evidence-based*

1b. Opportunity for Improvement: *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*

1c. Quality construct and rationale *(composite measures only)*

Subcriteron 1a: Evidence (page 32-38)

- Outcome measures
  - *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.*
- Structure, process, intermediate outcome measures
  - *The quantity, quality, and consistency of the body of evidence underlying the measure should demonstrate that the measure focuses on those aspects of care known to influence desired patient outcomes*
    - » Empirical studies (expert opinion is not evidence)
    - » Systematic review and grading of evidence
      - *Clinical Practice Guidelines – variable in approach to evidence review*
- For measures derived from patient (or family/parent/etc.) report
  - *Evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.*
  - *Current requirements for structure and process measures also apply to patient-reported structure/process measures.*

Criterion #1: Importance to measure and report

Criteria emphasis is different for new vs. maintenance measures

New measures	Maintenance measures
• Evidence – Quantity, quality, consistency (QQC) • Established link for process measures with outcomes	**DECREASED EMPHASIS**: Require measure developer to attest evidence is unchanged evidence from last evaluation; Standing Committee to affirm no change in evidence IF changes in evidence, the Committee will evaluate as for new measures
• Gap – opportunity for improvement, variation, quality of care across providers	**INCREASED EMPHASIS**: data on current performance, gap in care and variation

Criterion #2: Reliability and Validity–Scientific Acceptability of Measure Properties (page 40 -50)

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of health care delivery

2a. **Reliability (must-pass)**

2a1. Precise specifications including exclusions

2a2. Reliability testing—data elements or measure score

2b. **Validity (must-pass)**

2b1. Validity testing—data elements or measure score

2b2. Justification of exclusions—relates to evidence

2b3. Risk adjustment—typically for outcome/cost/resource use

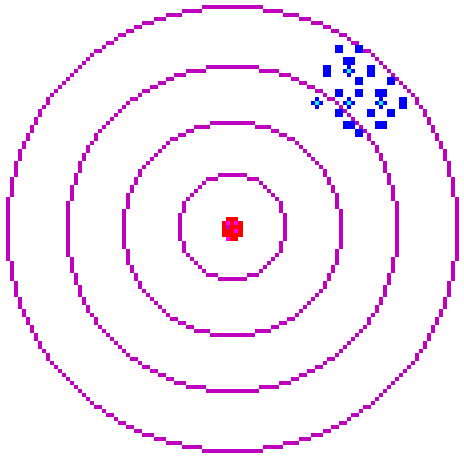
2b4. Identification of differences in performance

2b5. Comparability of data sources/methods

2b6. Missing data

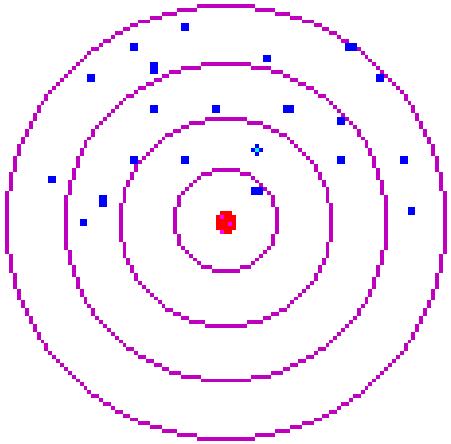
Reliability and Validity (page 41)

Assume the center of the target is the true score…



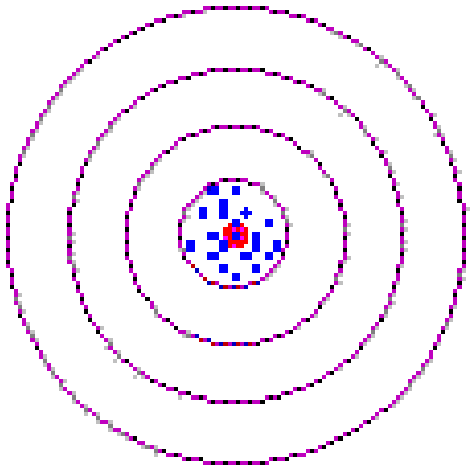
Reliable Not Valid

Consistent, but wrong



Neither Reliable Nor Valid

Inconsistent & wrong



Both Reliable And Valid

Consistent & correct

Evaluating Scientific Acceptability – Key Points (page 42)

Empirical analysis to demonstrate the reliability and validity of the *measure as specified,* including analysis of issues that pose threats to the validity of conclusions about quality of care such as exclusions, risk adjustment/stratification for outcome and resource use measures, methods to identify differences in performance, and comparability of data sources/methods.

Reliability Testing
Key points (page 43)

- Reliability of the **measure score** refers to the proportion of variation in the performance scores due to systematic differences across the measured entities in relation to random variation or noise (i.e., the precision of the measure).
  - *Example - Statistical analysis of sources of variation in performance measure scores (signal-to-noise analysis)*
- Reliability of the **data elements** refers to the repeatability/ reproducibility of the data and uses patient-level data
  - *Example – inter-rater reliability*
- Consider whether testing used an appropriate method and included adequate representation of providers and patients and whether results are within acceptable norms
- Algorithm #2

Validity testing (pages 45 - 49)

Empirical testing

- Measure score – assesses a hypothesized relationship of the measure results to some other concept; assesses the correctness of conclusions about quality
- Data element – assesses the correctness of the data elements compared to a "gold standard"

Face validity

- Subjective determination by experts that the measure appears to reflect quality of care
  - » Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.
  - » Requires systematic and transparent process, by identified experts, that explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Threats to Validity

- Conceptual
  - *Measure focus is not a relevant outcome of healthcare or not strongly linked to a relevant outcome*
- Unreliability
  - *Generally, an unreliable measure cannot be valid*
- Patients inappropriately excluded from measurement
- Differences in patient mix for outcome and resource use measures
- Measure scores that are generated with multiple data sources/methods
- Systematic missing or “incorrect” data (unintentional or intentional)

Criterion #2: Scientific Acceptability

New measures	Maintenance measures
• Measure specifications are precise with all information needed to implement the measure	NO DIFFERENCE: Require updated specifications
• Reliability • Validity (including risk-adjustment)	**DECREASED EMPHASIS**: If prior testing adequate, no need for additional testing at maintenance with certain exceptions (e.g., change in data source, level of analysis, or setting) Must address the questions regarding use of social risk factors in risk-adjustment approach

Criterion #3: Feasibility
(page 50-51)

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.

3a: Clinical data generated during care process

3b: Electronic sources

3c: Data collection strategy can be implemented

Criterion #4: Usability and Use (page 51-52)

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.

Use (4a) Must-pass for maintenance measures

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 six years after initial endorsement.*

4a2: Feedback by those being measured or others: *Those being measured have been given results and assistance in interpreting results; those being measured and others have been given opportunity for feedback; the feedback has been considered by developers.*

Usability (4b)

4b1: Improvement: *Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.*

4b2: Benefits outweigh the harms: *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).*

Criteria #3 & 4: Feasibility and Usability and Use

Feasibility

New measures	Maintenance measures
• Measure feasible, including eMeasure feasibility assessment	NO DIFFERENCE: Implementation issues may be more prominent

Usability and Use

New measures	Maintenance measures
• Use: used in accountability applications and public reporting	**INCREASED EMPHASIS**: Much greater focus on measure use and usefulness, including both impact and unintended consequences. Use is must pass sub-criterion.
• Usability: impact and unintended consequences	

Updated Guidance for Measures that Use ICD-10 Coding

- For CY2019 and beyond, reliability testing should be based on ICD-10 coded data.
- Validity testing should be based on ICD-10 coded data
- If providing face validity (FV), both FV of the ICD-10 coding scheme and FV of the measure score as an indicator of quality are a required update

eMeasures

- "Legacy" eMeasures
  - *Beginning September 30, 2017 all respecified measure submissions for use in federal programs will be required to the same evaluation criteria as respecified measures – the "BONNIE testing only" option will no longer meet endorsement criteria*

- For all eMeasures: Reliance on data from structured data fields is expected; otherwise, unstructured data must be shown to be both reliable and valid

NATIONAL QUALITY FORUM

Primary Care and Chronic Illness, Fall 2018 Measure Review Cycle

Standing Committee Measure Evaluation Web Meetings

Samuel Stolpe, PharmD, MPH
Suzanne Theberge, MPH
Hiral Dudhwala, RN, MSN/MPH

February 5, 2019

Agenda

- Welcome and Roll Call
- Consideration of Candidate Measures
- Public Comment
- Consideration of Related and Competing Measures
- Next Steps

Primary Care and Chronic Illness Fall 2018 Cycle Standing Committee

*New Committee Members

- **Dale Bratzler, DO, MPH (Co-Chair)**
- **Adam Thompson, BA (Co-Chair)**
- Thiru Annaswamy, MD
- Robert Bailey, MD
- Lindsay Botsford, MD
- Roger Chou, MD
- William Curry, MD, MS
- Jim Daniels, BSN
- Woody Eisenberg, MD*
- Kim Elliott, PhD
- V. Katherine Gray, PhD
- Starlin Haydon-Greatting, MS, BS, Pharm, FAPhA*
- Ann Kearns, MD, PhD
- Anne Leddy, MD, FACE
- Grace Lee, MD
- Anna McCollister-Slipp
- Janice Miller, DNP, CRNP, CDE
- James Rosenzweig, MD*
- Steven Strode, MD, Med, MPH, FAAFP
- William Taylor, MD
- Kimberly Templeton, MD
- John Ventura, DC

Welcome and Recap of Day 1

Consideration of Candidate Measures

Related and Competing Measures

Related and Competing Measures

If a measure meets the four criteria and there are endorsed/new related measures (same measure focus or same target population) or competing measures (both the same measure focus and same target population), the measures are compared to address harmonization and/or selection of the best measure.

Public Comment

Next Steps

Next Steps for Fall 2018 Cycle

- Draft Report Comment Period (30 days)
  - *March 18-April 16, 2019 (tentative)*
- Committee Post-Comment Web Meeting
  - *May 6, 2019 2-4 pm EST*

Spring 2019 Cycle Updates

- Intent to submit deadline – January 7, 2019
- 12 measures submitted
  - *11 maintenance measures*
  - *1 new measure*
  - *3 measures considered "complex" and will be reviewed by NQF Scientific Methods Panel for scientific acceptability criterion.*

Project Contact Info

- Email: primarycare@qualityforum.org
- NQF phone: 202-783-1300
- Project page: http://www.qualityforum.org/Primary_Care_and_Chronic_Illness.aspx
- SharePoint site: http://share.qualityforum.org/Projects/Primary%20Care%20and%20Chronic%20Illness/SitePages/Home.aspx

Questions?

THANK YOU

Adjourn