NATIONAL QUALITY FORUM
Driving measurable health improvements together

http://www.qualityforum.org

Scientific Methods Panel Web Meeting

Trina Calloway-Peden
Mike DiVecchia
Hannah Ingber
Caitlin Flouton

May 26, 2020

Welcome & Introductions

NATIONAL QUALITY FORUM

NQF Scientific Methods Panel Team

- Senior Leads
  - Trina Calloway-Peden, MS
- Project Management
  - Mike DiVecchia, PMP
  - Hannah Ingber, MPH
  - Caitlin Flouton, MS

Scientific Methods Panel Members

J. Matt Austin, PhD	Jack Needleman, PhD
Bijan Borah, MSc, PhD	David Nerenz, PhD, Co-chair
John Bott, MBA, MSSW	Eugene Nuccio, PhD
David Cella, PhD, Co-chair	Sean O'Brien, PhD
Daniel Deutscher, PT, PhD	Jennifer Perloff, PhD
Lacy Fabian, PhD	Patrick Romano, MD, MPH
Marybeth Farquhar, PhD, MSN, RN	Sam Simon, PhD
Jeffrey Geppert, EdM, JD	Alex Sox-Harris, PhD, MS
Laurent Glance, MD	Michael Stoto, PhD
Joseph Hyder, MD	Christie Teigland, PhD
Sherrie Kaplan, PhD, MPH	Ronald Walters, MD, MBA, MHA, MS
Joseph Kunisch, PhD, RN-BC, CPHQ	Terri Warholak, PhD, RPh, CPHQ, FAPhA
Paul Kurlansky, MD	Eric Weinhandl, PhD, MS
Zhenqiu Lin, PhD	Susan White, PhD, RHIA, CHDA

Meeting Overview

NATIONAL QUALITY FORUM

Meeting Agenda

- Discuss Scientific Acceptability of Measure Properties: **Reliability**
  - Does the term "critical data element" include data elements used in risk-adjustment models? All data elements in risk-adjustment models?
  - Is it ever acceptable to have just data element reliability? If not, should the current requirement for either data element OR measure score reliability be changed?
  - Does data element validity guarantee data element reliability, so that showing validity removes the requirement to show reliability?
- Discuss Scientific Acceptability of Measure Properties: **Validity**
  - Should face validity continue to be accepted as the minimum requirement for new measure submissions?
  - Should we require both data element and measure score validity testing?
  - What guidance can we offer developers about the range of variables that can be used to establish validity by their correlation with the specific measure being evaluated?
- Public Comment
- Next Steps

Scientific Acceptability of Measure Properties – Reliability

Review of Existing NQF guidance

Criterion #2: Reliability–Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) results about the quality of healthcare delivery

2a. Reliability (must-pass)

- 2a1. Precise specifications including exclusions
- 2a2. Reliability testing—data elements or measure score

NATIONAL QUALITY FORUM

Empirical Testing

- The approach to empirical testing as it relates to the measure construct of administrative claims compared to another measure with the same data elements
- Administrative claims also used as performance scores
- SMP - is there a concern with the current method that developers are using which does not represent correlations to an independent variable or measure?

NATIONAL QUALITY FORUM

NQF Definitions of Reliability

- Repeatability (consistency, reproducibility, stability)
- Precision
- Data Element Reliability: Repeatability and reproducibility of the data elements for the same population in the same time period
- Measure Score Reliability: Precision: Proportion of variation in the performance scores due to systematic differences across the measured entities (signal) in relation to random error (noise)

NATIONAL QUALITY FORUM

Current Assumptions about Reliability

- There can be minor error in performance measurement
  - Random error affects reliability; systematic error affects validity
- Reliability is **not a static property of a measure** (it can vary under conditions of implementation)
- Reliability is not an all-or-none property and is instead a matter of degree
  - Considerations are scope of testing, method used, and results obtained
- Reliability does not guarantee validity

NATIONAL QUALITY FORUM

Reliability Testing – Data Element

- Reliability of the **data elements** refers to the repeatability/ reproducibility of the data for the same population in the same time period
  - Common Approaches
    - » inter-rater/abstractor or intra-rater/abstractor agreement
    - » **internal consistency for multi-item scales**
    - » **test-retest for surveys or other forms of subjective measurement**
- Current NQF Guidance
  - All critical data elements must 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.

NATIONAL QUALITY FORUM

Current Testing Requirements

	Structure/process/outcome	PRO-PM / Instrument Based	Composite	eCQM
Reliability	Element OR score ("short-cut"* allowed)	Element AND score	Score	Depends on how data are stored
Validity	Element OR score or face validity**	Element AND score	New: element OR score OR face validity Maintenance: Score	Element

*No reliability testing required if data element validity testing conducted and results are adequate

** Face validity allowed for new measures, but only with accepted justification at maintenance

Scientific Acceptability of Measure Properties – Validity

Review of Existing NQF Guidance

Criterion #2: Validity – Scientific Acceptability of Measure Properties

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

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

NATIONAL QUALITY FORUM

Empirical Validity Testing — Measure Score

Challenging Examples

- Comparing CAHPS measures to themselves
- Behavioral health (substance use disorder (SUD) screening versus depression and infectious disease screening) versus actual better SUD outcomes
- Cost measures comparing to other claims-based measures with the same data elements, construct vs. content validity
  - Considerations from the Cost and Efficiency Standing Committee

NATIONAL QUALITY FORUM

Additional Context

NQF currently does not require validity testing relative to:

- An expected outcome (e.g., process measure about foot exams for patients with diabetes does not have to be correlated to a measure about foot amputation)
- Testing is not limited to other NQF-endorsed measures
- Testing does not have to use an "external" measure or dataset
  - e.g., we allow testing of an instrument-based domain measure (e.g., treated with respect") with a "global" measure from the same instrument (e.g., would you recommend this agency)
  - Recently, some concerns about "circular" testing (e.g., stability over time)

NATIONAL QUALITY FORUM

Panel Considerations

- Face Validity
  - Should face validity continue to be accepted as the minimum requirement for new measure submissions? If so, what guidance should we offer beyond that already in place about what the criteria for acceptable face validity testing should be?
- Empirical Validity Testing
  - Should we require both data element and measure score validity testing?
  - What guidance can we offer developers about the range of variables that can be used to establish validity by their correlation with the specific measure being evaluated?
    - » Is it acceptable, for example, to validate a measure of screening for something by correlating it with a measure of screening for something else?
    - » What level of correlation is acceptable to establish validity? Is any statistically significant correlation acceptable?

NATIONAL QUALITY FORUM

Tentative Next Steps for Criteria Recommendations

- May 2020: Obtain consensus recommendations from SMP during monthly call
- NQF consideration of recommendations
- Public Commenting
- November 2020: Present SMP recommendations to CSAC
  - CSAC may accept/reject/modify the recommendations
  - CSAC may suggest an implementation timeframe
- Winter/Early spring: Begin to publicize changes to criteria
- NOTE that NQF often allows up to a 1-year gap between changing criteria and implementing the changes
  - Likely, any SMP-recommended changes would not be required of developers until August 2021 (or even Spring Cycle 2022)

Opportunity for Public Comment

Next Steps

NATIONAL QUALITY FORUM

White Papers

Published

- NQF Guidelines for Evaluating the Scientific Acceptability of Risk-Adjusted Clinical Outcome Measures (Larry G. et al.)
- The NQF Scientific Methods Panel (David N. et al.)

Next

- Reliability
  - Led by Dave N
- Social Risk Adjustment
  - Led by Dave N
- Evaluation Cost/Resource Use Measures
  - Led by Jenn, Jack, and Bijan

NATIONAL QUALITY FORUM

Next Steps

- Next Intent to Submit deadline for Fall 2020: August 3, 2020

NATIONAL QUALITY FORUM

2020 SMP Meetings

Meeting Date	Tentative Topic/Activity
May 26 - 1-3PM ET	Wrap up Criteria Recommendations
June 16 - 2-4PM ET	Continue reliability guidance discussion
July 21 - 2-4PM ET	Validity Testing Guidance: Choosing a comparator
August 25 - 1-3PM ET	TBD
October 28-29 - all-day, in-person meeting	Measure Evaluation
December 8 - 1-3PM ET	TBD

NATIONAL QUALITY FORUM

Project Contact Info

- Email: MethodsPanel@qualityforum.org
- NQF phone: 202-783-1300
- Project page: http://www.qualityforum.org/Measuring_Performance/Scientific_Methods_Panel.aspx
- SharePoint site: http://share.qualityforum.org/Projects/NQF%20Scientific%20Methods%20Panel/SitePages/Home.aspx

THANK YOU.

NATIONAL QUALITY FORUM

http://www.qualityforum.org