



**NATIONAL
QUALITY FORUM**

Driving measurable health
improvements together

<https://www.qualityforum.org>

Scientific Methods Panel Web Meeting

December 14, 2021

*Funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I -
HHSM-500-T0001.*

Housekeeping Reminders

- This is a Webex meeting with audio and video capabilities:
 - ▣ **Meeting link:** <https://nqf.webex.com/nqf/j.php?MTID=mf1060e2b56732d1cfb9aaaf89d6863ab>
 - ▣ **Meeting number:** 2347 337 9032
 - ▣ **Password:** MScAEvent
 - ▣ Optional: Dial 1-844-621-3956 and enter passcode 2347 337 9032
 - ▣ Please place yourself on mute when you are not speaking
- We encourage you to use the following features
 - ▣ Chat box: to message NQF staff or the group
 - ▣ Raise hand: to be called upon to speak
- We will conduct Scientific Methods Panel roll call once the meeting begins

If you are experiencing technical issues, please contact the NQF project team at methodspanel@qualityforum.org

Welcome

NQF Scientific Methods Panel Team

- Tricia Elliott, MBA, CPHQ, FNAHQ, Senior Managing Director
- Mike DiVecchia, MBA, PMP, Senior Project Manager
- Hannah Ingber, MPH, Senior Analyst
- Gabby Kyle-Lion, MPH, Coordinator
- Sharon Hibay, DNP, RN, Senior Consultant
- Jill Boylston Herndon, PhD, Consultant
- Elisa Munthali, MPH, Consultant

Scientific Methods Panel Members

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

Meeting Objectives

- Discuss feedback on minimum reliability testing thresholds from the Consensus Standards Approval Committee (CSAC), the measure developer advisory panel, and the Standing Committee advisory group
- Confirm consensus on recommendations of minimum reliability testing thresholds for patient/encounter and accountable entity levels
- Discuss feedback on accountable entity level reliability and validity testing requirements from the CSAC, the measure developer advisory panel, and the Standing Committee advisory group
- Confirm consensus on recommendations for reliability and validity testing at the accountable entity level for maintenance measures

Meeting Agenda

- Confirm consensus on recommendations for minimum reliability testing thresholds
- Confirm consensus on recommendations for accountable entity level reliability and validity testing requirements for maintenance measures
- Opportunity for public comment
- Wrap Up/Next steps

Minimum Reliability Testing Thresholds

Reliability Testing Thresholds - Questions from the CSAC

- During the October 12th meeting, the CSAC reviewed the minimum reliability testing thresholds for the *recommended* accountable entity level and *proposed* patient/encounter level tables
- The CSAC expressed widespread appreciation to SMP for this work and stated this guidance will reduce measure testing burdens for developers, implementers, and measure evaluators. They favored the direction of the evolving recommendations and wish to stay informed of the process, progress, and findings of the SMP and from measurement community feedback
- For patient/encounter level reliability testing, CSAC members suggested that Kappa minimums be higher than 0.4, as many NQF endorsed measures are used for accountability purposes
- The CSAC made recommendations for a final guidance document, including:
 - ▣ Guiding principles for the threshold application (e.g., SMP and Standing Committees will use the thresholds to justify measure ratings)
 - ▣ Next steps for measures that do/do not meet testing minimums (e.g., Measures will not be automatically rejected by NQF Staff based on testing results, and Standing Committees may pull eligible SMP reviewed measures for discussion and potential revote)
 - ▣ Rationale and cited references for each testing threshold recommendation

Reliability Testing Thresholds - Questions from the Measure Developer Advisory Panel

- The Measure Developer Advisory Panel members reviewed the *recommended* accountable entity level and *proposed* patient/encounter level table during the November 8, 2021 meeting and agreed with the general direction, framework, and proposed thresholds.
- Members requested greater detail and guidance on testing requirements and methods for:
 - ▣ Testing thresholds specific to the data collection method (i.e., instruments and surveys for PROMs) versus measure type (i.e., PRO-PMs)
 - ▣ PRO-PMs versus clinician-reported/-interpreted measures
 - ▣ Increasing patient/encounter Kappa thresholds above 0.4
- Members asked for clarification on potential exceptions to the thresholds. NQF stated they would be evaluated for well-documented and strong rationale or justification, such as:
 - ▣ Measures that target low-volume/high-priority “never” events
 - ▣ Supporting literature explaining low testing results
 - ▣ Other testing requirement constraints by measure or data collection types (e.g., eCQMs, sample sizes, level of analysis)

Reliability Testing Thresholds - Questions from the Standing Committee Advisory Panel

- The Standing Committee Advisory Panel reviewed the *recommended* accountable entity level and *proposed* patient/encounter level table during the November 18, 2021 meeting and agreed with the general direction, framework, and proposed thresholds.
- Members agreed the table and a well-planned roll-out are vital for consistent CDP Standing Committee measure evaluation
- Members recommended that successful implementation should include robust documented guidance and training for all stakeholders, including:
 - ▣ The purpose and use of the threshold tables to serve as a discussion guide for Standing Committee evaluations
 - ▣ SMP's decision-making framework, supporting references, refuting rationales, and justifications for not meeting the thresholds
 - ▣ A supplement of graduated testing ranges below, above, and near the thresholds to guide interpretation of results.
 - ▣ Members also agreed the 0.4 patient/encounter Kappa threshold is too low.
 - ▣ Members also asked NQF staff to analyze the impact of these decisions on current measure portfolios.¹¹

Person-/Encounter-Level Reliability Testing (i.e., data element testing)

Approach (Test)	Purpose	Range	Threshold
Internal consistency (e.g., Cronbach's Alpha)	The internally consistency of items in a multi-item scale.	0 to 1	0.7
Inter-rater agreement e.g., (Cohen's Kappa)	The inter-rater agreement of qualitative items correcting for chance.	-1 to +1	0.4
Test-Retest Reliability (Intraclass coefficient [ICC] or Pearson correlation)	Extent to which two measurements of the same concept at different times agree.	-1 to +1	0.5
Linear Relationships (e.g., Pearson correlation coefficient)	Agreement between two measures of the same concept.	-1 to +1	0.6

Accountable Reporting Entity Level Reliability Testing (i.e., performance measure score testing)

Approach (Test)	Testing Purpose	Range	Threshold
Signal to Noise Ratio (SNR) or Inter-Unit Reliability (IUR)	The precision attributed to an actual construct versus random variation.	0 to 1	0.6
Split-half reliability (Intraclass coefficient, with correction for full sample with Spearman-Brown formula)	Agreement between two measures of the same concept derived from split samples drawn from the same entity at a single point in time.	0 to 1	0.6

Recommendations for Reliability Thresholds Timeline



Date	Proposed Activity
10/12/2021	CSAC Input garnered
10/27/2021	SMP Reliability Discussion (Deferred to 12/14)
11/8/2021	Measure Developer Advisory Panel Review
11/18/2021	Standing Committee Advisory Group Review
12/14/2021	SMP Review/Status/Confirm consensus
Q1 2022	NQF leadership Review (CSO)
Q1 2022	Public Comment
Q2 2022	CSAC Review/Status/Approval
Q3 2022	Implementation in Guidance documents

Accountable Entity Level Reliability and Validity Testing Requirements for Maintenance Measures

Reliability Testing at the Accountable Entity Level for Maintenance Measures – CSAC, Measure Developer, and Standing Committee Advisory Panel Feedback

- Requiring accountable entity-level reliability testing for all maintenance measures tests the consistency of specifications implemented in practice.
- Members agreed with the recommendation and asked for clarification for:
 - ▣ Data collection from measure implementation may take longer than the current maintenance cycle of three years
 - ▣ Rare events / “never” events that often yield low data volumes
 - ▣ Measure and reliability results that require broader Standing Committee discussion to evaluate all criteria
- NQF will require a strong justification/explanation for SMP and Standing Committee evaluation when this requirement cannot be met.

Validity Testing at the Accountable Entity Level for Maintenance Measures - CSAC, Measure Developer, and Standing Committee Advisory Panel Feedback

In July 2021, the SMP made two recommendations for validity testing in maintenance evaluations:

1. Validity testing at the accountable-entity level should be required for all maintenance measures.
 - *If measure developers are unable to meet this requirement, NQF should require, and developers should provide, a strong rationale supporting this rare instance.*
2. Measures submitted for maintenance evaluations with face validity testing should also include empirical validity testing or, in rare instances, patient-/encounter-level validity testing with a strong rationale for not performing accountable-entity validity testing
 - Members largely agreed, but raised additional considerations:
 - ▣ What rationale(s) would justify a lack of accountable entity validity testing at maintenance?
 - ▣ What rationale(s) would justify a lack of accountable entity empirical validity testing at maintenance, such as no a suitable comparator for testing?
 - ▣ NQF should conduct an analysis of the impact of these recommendations on current measure portfolios.

Evaluating Testing at the Accountable Entity Level for Maintenance Measures

At past meetings, SMP members also agreed that further discussion was needed regarding the following item:

- Prioritizing accountable entity-level testing when patient/encounter-level testing is also present. NQF's current policy allows for measures with moderate patient/encounter-level testing to pass, despite having poor accountable entity-level testing.

Note: Although many measures in NQF's portfolio are used for accountability purposes, no priority is set on the results of accountable entity level testing over patient/encounter level testing. Additionally, current evaluation criteria guides evaluators to consider all submitted testing in criterion rating, even if the provided testing is not required (e.g., accountable entity and patient/encounter reliability and validity testing conducted for a new process measure)

Question for the SMP consideration:

- How would the SMP consider prioritizing results for accountable entity level testing and patient/encounter-level testing for reliability and validity?

Requiring Reliability and Validity Testing at the Accountable Entity Level for Maintenance Measures – Timeline

Date	Proposed Activity
10/12/2021	CSAC Input garnered
10/27/2021	SMP Reliability Discussion (Deferred to 12/14)
11/8/2021	Measure Developer Advisory Panel Review
11/18/2021	Standing Committee Advisory Group Review
Today! → 12/14/2021	SMP Review/Status/Confirm consensus
Q1 2022	NQF leadership Review (CSO)
Q1 2022	Public Comment
Q2 2022	CSAC Review/Status/Approval
Q3 2022	Implementation in Guidance documents

Potential Items for Future SMP Discussion

- Face validity testing requirements and acceptability
- Formative vs. reflective composite models
- Incorporating intended use into scientific acceptability discussions
- High, moderate, low, and insufficient ratings and their use
- Further clarification on divergent testing results at the patient/encounter and accountable entity levels
- Appropriate testing sample size requirements

Opportunity for Public Comment

Next Steps

Next Steps and Reminders

- Meeting summary review
- NQF staff to conduct impact analyses on measure portfolio for the following recommendations:
 - ▣ Minimum reliability testing thresholds
 - ▣ Maintenance reliability and validity testing requirements
- SMP Advisory Meetings (every other month between SMP Evaluation Meetings)
 - ▣ Next SMP Advisory Meeting to be held in February 2022 (specific date/time TBD)
- SMP Evaluation Meetings are planned for:
 - ▣ Spring 2022 in late March 2022
 - ▣ Fall 2022 in late October 2022

Thank You, Lacy Fabian!

- Lacy Fabian has been on the Scientific Methods Panel since September 2017
- We are grateful for her expertise and contributions to this important work
- On behalf of the entire team at NQF, thank you!

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:
<https://share.qualityforum.org/portfolio/ScientificMethodsPanel/SitePages/Home.aspx>

Adjourn

THANK YOU.

NATIONAL QUALITY FORUM

<https://www.qualityforum.org>