



Scientific Methods Panel Monthly Call Meeting

Karen Johnson, MS
Andrew Lyzenga, MPP
Poonam Bal, MHSA
May Nacion, MPH
Miranda Kuwahara, MPH

July 12, 2018

Welcome, Roll Call, and Review of Meeting Objectives

Scientific Methods Panel Members

- David Cella, PhD, (Co-Chair)
- David Nerenz, PhD (Co-Chair)
- J. Matt Austin, PhD
- Bijan Borah, MSc, PhD
- John Bott, MBA, MSSW
- Lacy Fabian, PhD
- Marybeth Farquhar, PhD, MSN, RN
- Jeffrey Geppert, EdM, JD
- Paul Gerrard, BS, MD
- Laurent Glance, MD
- Stephen Horner, RN, BSN, MBA
- Karen Joynt Maddox, MD, MPH

Scientific Methods Panel Members (continued)

- Sherrie Kaplan, PhD, MPH
- Joseph Kunisch, PhD, RN-BC, CPHQ
- Paul Kurlansky, MD
- Zhenqiu Lin, PhD
- Jack Needleman, PhD
- Eugene Nuccio, PhD
- Jennifer Perloff, PhD
- Sam Simon, PhD
- Michael Stoto, PhD
- Christie Teigland, PhD
- Ronald Walters, MD, MBA, MHA, MS
- Susan White, PhD, RHIA, CHDA

Meeting Objectives



Process Updates

Methodologic Discussion: Reliability for
Instrument-Based Performance Measures

Methods Panel Process Updates

Panel Recommendations to Improve Process: In-Person Meeting

- The process: Suggested “Option 1.5” (something between current independent evaluation and full Panel evaluation of all measures)
 - *Learning from each other is a key “ask”*
 - *Consensus from a larger group would be better*
 - *HOWEVER, impossible to evaluate up to 50 measures in a 2-3 day in-person meeting*
 - » Might work if they discuss only those measures where there is disagreement
 - » BUT there are “cons” to doing it this way
 - *Regardless, agreed we need earlier resolution*
 - *Wanted more info from us about final ratings, points of disagreement*
- The evaluation form: Mostly free-text but some checkboxes

New Process: Most Likely Moving to Subgroup Evaluations

- Subset of panel members evaluate a subset of measures
- Requires series of webinars
 - *Mandatory discussion of measures where panelists disagree on ratings; can discuss other measures if desired*
 - *We will assign discussants for measures*
- Ratings based on **subgroup vote** (no co-chair adjudication required)
- Webinars will be open to the public (but won't allow developer input)
- Assigned workgroups will submit pre-meeting comments in lieu of completing current evaluation form
- Output is a simple majority vote and staff-generated summary + pre-meeting comments

“Toolkit” / White Paper Progress to Date

- What we need to do and want to do
 - *Create a “methods toolkit” beneficial to all (staff/panel members/public/developers) that outlines definitions, methodology, thresholds, and best practices*
 - *Submit to peer-reviewed journals key methodologic information based on Methods Panel discussions (“subset” of the toolkit, but different format)*
 - » Three papers conceptualized: overview of MP work; clinical outcomes (focus on risk-adjustment; reliability/validity discussions included within); patient-reported outcomes
- Writing workgroups will take responsibility for majority of writing
 - *Leads identified (David, Dave, Larry, probably NQF-Karen)*
 - *Volunteer if interested and available*
 - *Target is good draft by end of summer for 1st two papers*

Methodologic Discussion: Reliability for Instrument-Based Performance Measures

Expectations to date

- For instrument-based measures, reliability should be demonstrated for both the data (instrument) and the performance measure score
 - *Data element reliability focuses on the question(s)/item(s) in the instrument*
 - » Specific types of analysis dependent on instrument structure and how the performance measure is constructed
 - *Score-level reliability focuses on ability to distinguish differences in performance between providers (data aggregated at the provider level)*
 - » Assumption is that if there are not meaningful differences between providers, NQF would not endorse the measure

Questions to Consider

- For performance measures **based on multi-item scales**, is there an “in-between” analysis that we should expect (i.e., another “level” of testing)?
 - *To show the reliability of the patient-level scores that result when combining the multiple items*
- If so:
 - *What kind of analysis would that entail?*
 - *Do we still need to know about the reliability of the individual items?*
 - *What language should we use to describe?*

Member and Public Comment

Next Steps

- Monthly 1 hour calls
 - *Every 2nd Thursday of the month*
 - *Next call: August 9, 2018 3pm ET*
- Contact Information: methodspanel@qualityforum.org

Adjourn