



NATIONAL
QUALITY FORUM

Scientific Methods Panel Monthly Call Meeting

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

January 11, 2018

Welcome, Roll Call, and Review of Meeting Objectives

Scientific Methods Panel Members

Co-Chairs

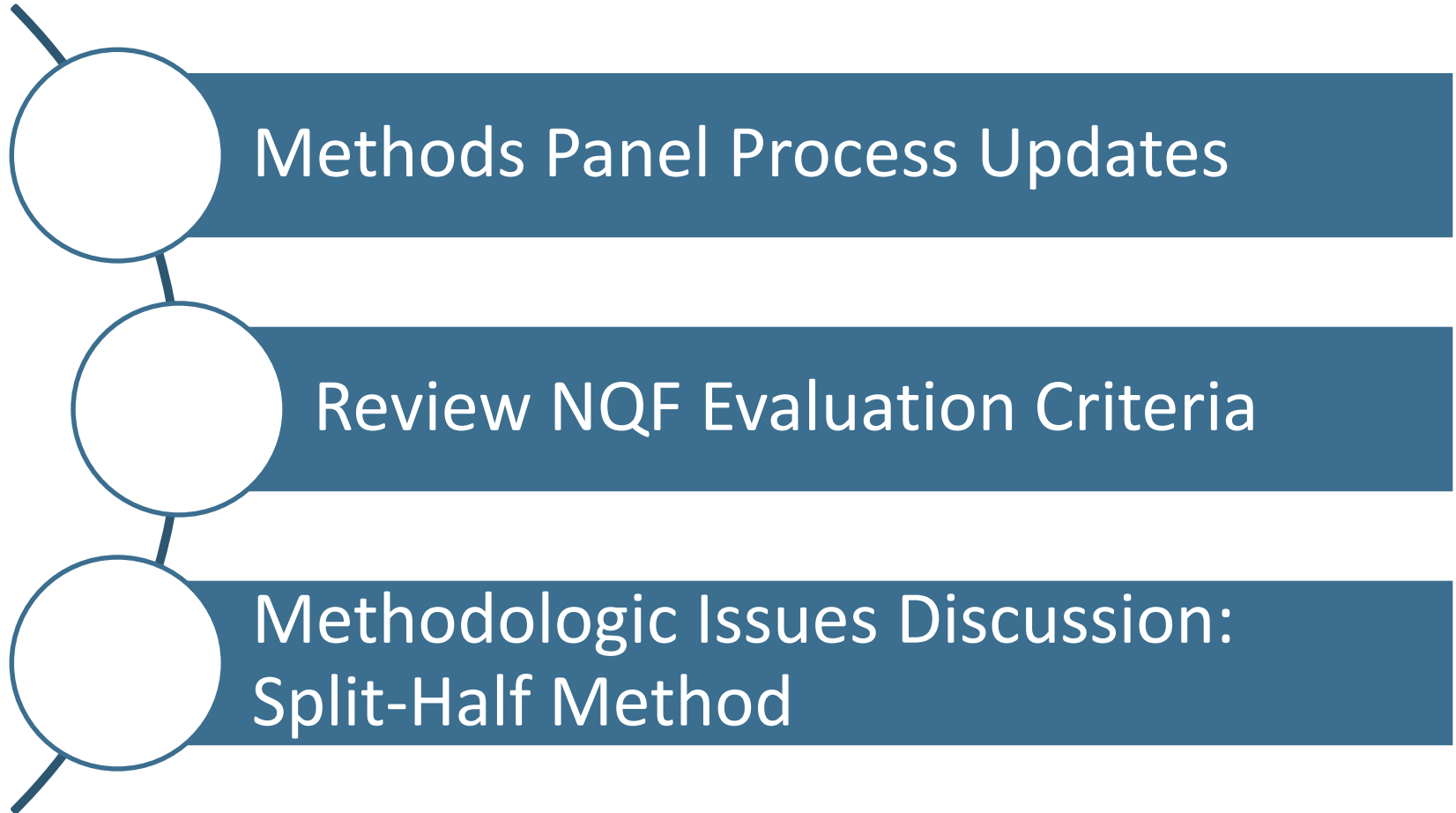
David Cella, PhD

Karen Joynt Maddox, MD, MPH

Panel Members

J. Matt Austin, PhD	Paul Kurlansky, MD
Bijan Borah, MSc, PhD	Zhenqiu Lin, PhD
John Bott, MBA, MSSW	Jack Needleman, PhD
Lacy Fabian, PhD	David Nerenz, PhD
Marybeth Farquhar, PhD, MSN, RN	Eugene Nuccio, PhD
Jeffrey Geppert, EdM, JD	Jennifer Perloff, PhD
Paul Gerrard, BS, MD	Sam Simon, PhD
Laurent Glance, MD	Michael Stoto, PhD
Stephen Horner, RN, BSN, MBA	Christie Teigland, PhD
Sherrie Kaplan, PhD, MPH	Ronald Walters, MD, MBA, MHA, MS
Joseph Kunisch, PhD, RN-BC, CPHQ	Susan White, PhD, RHIA, CHDA

Meeting Objectives



Methods Panel Updates

Process Updates

- Andrew Lyzenga: Newest NQF staff member on our Methods Panel Team
- Will allow informal discussions between evaluators (phone or e-mail) but still require separate evaluations
- Discussion board implemented
 - *SharePoint now includes a discussion board*
 - » For discussing methodologic topics, *not* specific measures
- Subgroups to discuss methodologic topics
 - *Risk-adjustment*
 - *Testing methodologies*
 - *Emerging measurement approaches (e.g., instrument-based measures, eMeasures)*

Process Updates

- Preliminary Analysis [Evaluation] Form
 - *Revisions*
 - » Revised directions
 - » Continuous numbering
 - » Reordering questions
 - *Quality Assurance process*
 - » Combination of phone calls and email between panel members and NQF staff regarding any issues with completing the PA Form
 - *Reminder: Your responses will be made publicly available*

Understanding the NQF Evaluation Criteria

Current Evaluation Criteria

- Resources available to you
 - *NQF's Measure Evaluation Criteria document and the "key points" document*
 - *Staff: e-mail or phone calls*
 - *Clarifications from developers*
 - » One-time option
 - » Must be very early on in the process
- eMeasures (eCQMs):
 - *Testing from 2 EHR systems required*
 - » While more would be great, it is not required
 - *Reliability testing not required if based on data from structured data fields. Unstructured fields require both reliability and validity testing*
 - *Will also provide feasibility scorecard to you*

Current Evaluation Criteria (continued)

- Instrument-based measures
 - *For reliability and validity, require testing at **both** levels*
 - » Data element level → must demonstrate R/V of the instrument
 - » Measure score level → testing of the actual performance measure
 - *We do allow multiple performance measures under same NQF number: need only one form, but may require multiple ratings*
- Why did we ask about data element validity in the Reliability section?
 - *If data element validity testing provided, we **do not** require additional reliability testing – but we still need a rating for reliability*

Current Evaluation Criteria (continued)

- Why did we (often) ask you to skip the face validity questions?
 - *Face validity “testing” means something specific to NQF*
 - » Can scores be used to differentiate good from poor quality?
 - » We see content validity of instruments as something different
 - *FV is acceptable for new measures, but empirical testing is expected for measures up for re-endorsement, unless there is good justification for lack of empirical testing*
 - » FV is the weakest form of validation—so ratings should hinge on empirical results if they are available
- Risk-adjustment: Why did we frown upon responses of “Not Applicable”?
 - *NQF typically expects risk-adjustment of complex measures*
 - *If not risk-adjusted, are you convinced it isn’t needed?*

Questions?

Methodologic Issue: Split-Half Method for Score-Level Testing

Background

- Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
 - *Is the variation between providers primarily due to real differences? Or is it because there is a lot of "noise" in the measurement?*

- “Split-Half” (sometimes called “test-retest”) Method:
 - *Randomly split each hospital’s cases into two subsets*
 - *Compute measure for each subset for each hospital?*
 - *Calculate ICC [2,1] between the subsets across the hospitals*
 - » Sometimes, bootstrapping used
 - » Maybe Pearson’s correlation rather than ICC (?) [staff will clarify]

Questions to Consider

- Can this method answer the question of ability to distinguish between providers?
- Should we continue to accept this method of testing?
- Any recommendations regarding this testing? (e.g., nomenclature, statistics, etc.)?

Upcoming Methodologic Discussions

- Can we determine thresholds (or rules of thumb_ for reliability?
- What are the various methods that can be used to demonstrate reliability?
- Other??

Member and Public Comment

Next Steps

- Monthly 1 hour Calls
 - *Every 2nd Thursday of the month*
 - *Next call: February 8, 3pm ET*
- Complete Disclosure of Interest survey for Spring Cycle 2018 measures by **January 19**
- Complete measure evaluations by **February 20**
- In-Person meeting: Likely in May 2018
- Contact Information: methodspanel@qualityforum.org

THANK YOU