NATIONAL QUALITY FORUM

National Consensus Standards for Renal Conditions

Standing Committee Orientation

Andrew Lyzenga, MPP
Poonam Bal, MHSA

May 3, 2018

Welcome

Project Team

Andrew Lyzenga, MPP
Senior Director

Poonam Bal, MHSA
Senior Project Manager

Agenda for the Call

- Standing Committee Introductions
- Review of project activities and timelines
- Overview of NQF's measure evaluation criteria
- SharePoint Tutorial
- Measure Worksheet example
- Next steps

Renal Standing Committee

- Constance Anderson, BSN, MBA (Co-Chair)
- Lorien Dalrymple, MD, MPH (Co-Chair)
- Ishir Bhan, MD, MPH
- Rajesh Davda, MD, MBA, CPE
- Elizabeth Evans, DNP
- Michael Fischer, MD, MSPH
- Renee Garrick, MD, FACP
- Stuart Greenstein, MD
- Mike Guffy
- Debra Hain, PhD, APRN, ANP-BC, GNP-BC, FAANP
- Lori Hartwell
- Frederick Kaskel, MD, PhD
- Myra Kleinpeter, MD, MPH
- Alan Kliger, MD
- Mahesh Krishnan, MD, MPH, MBA, FASN
- Lisa Latts, MD, MSPH, MBA, FACP
- Karilynne Lenning, MHA, LBSW
- Franklin Maddux, MD, FACP
- Andrew Narva, MD, FACP, FASN
- Jessie Pavlinac, MS, RD, CSR, LD
- Mark Rutkowski, MD
- Michael Somers, MD
- Bobbi Wager, MSN, RN
- John Wagner, MD, MBA
- Joshua Zaritsky, MD, PhD

Measure Evaluation Criteria Overview

NQF Measure Evaluation Criteria for Endorsement

NQF endorses measures for accountability applications (public reporting, payment programs, accreditation, etc.) as well as quality improvement.

- Standardized evaluation criteria
- Criteria have evolved over time in response to stakeholder feedback
- The quality measurement enterprise is constantly growing and evolving – greater experience, lessons learned, expanding demands for measures – the criteria evolve to reflect the ongoing needs of stakeholders

Major Endorsement Criteria *(page 28)*

- ***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

Criterion #1: Importance to Measure and Report (page 30-39)

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

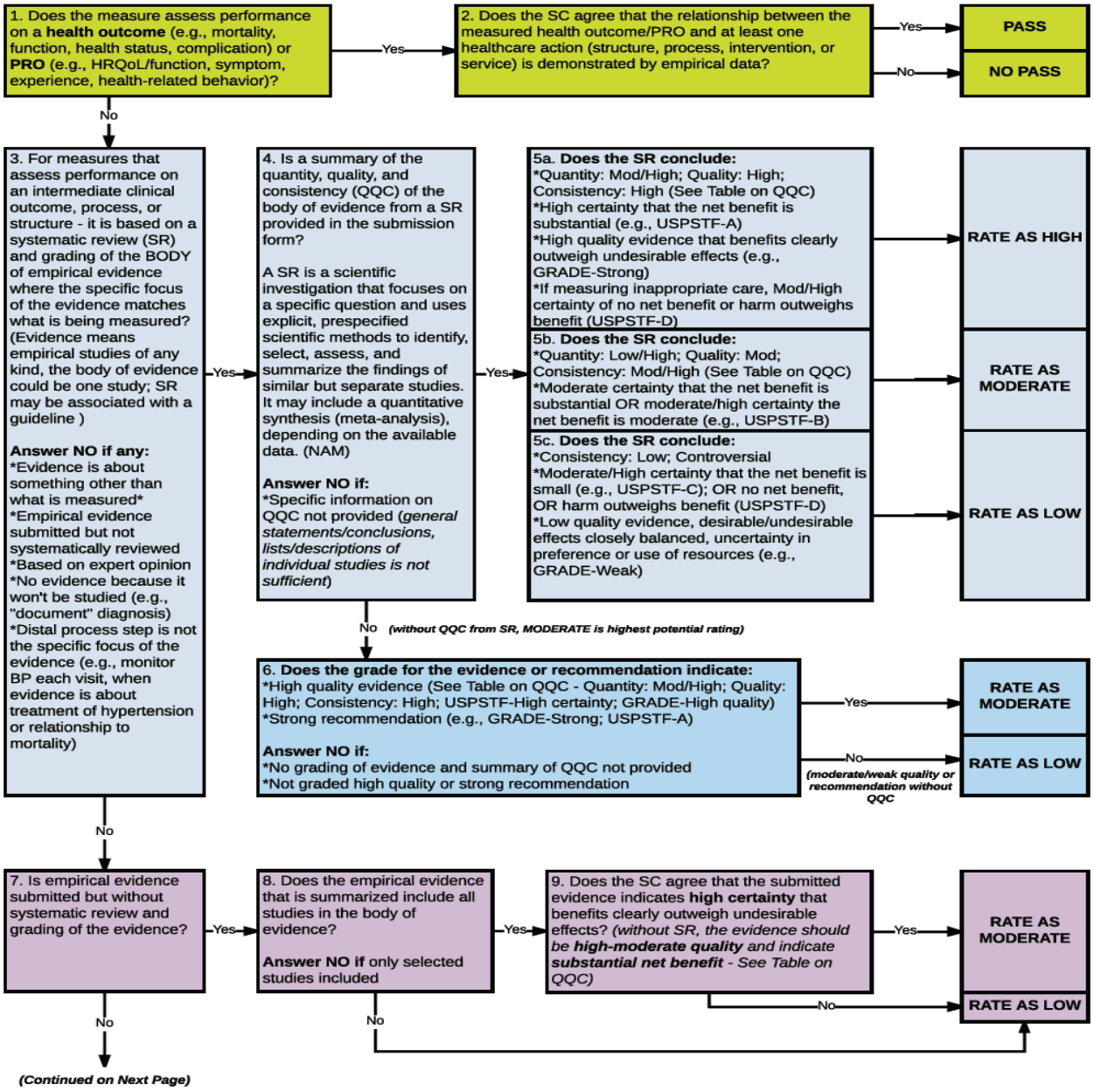
   1b. Opportunity for Improvement: demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or disparities in care across population groups

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

Subcriteron 1a: Evidence (page 31-37)

- **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.*

Rating Evidence: Algorithm #1 – page 43



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 from last evaluation; Standing Committee to affirm no change in evidence IF evidence has changed, 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 39 - 48)

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

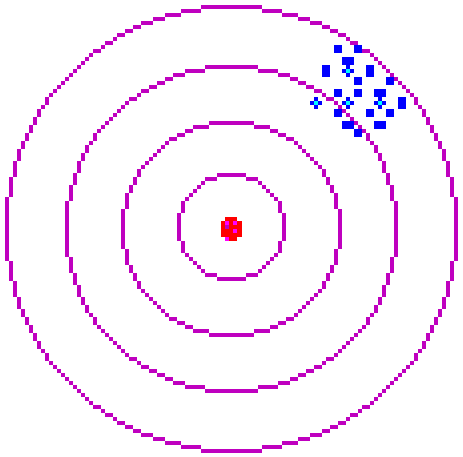
2b4. Identification of differences in performance

2b5. Comparability of data sources/methods

2b6. Missing data

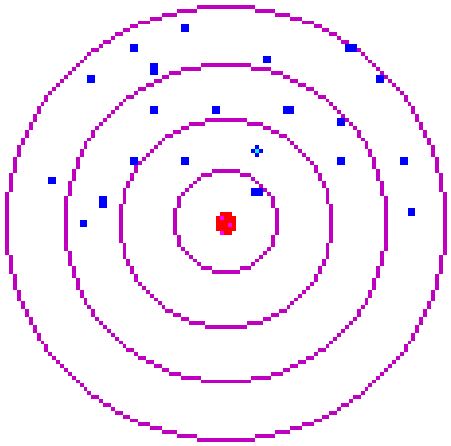
Reliability and Validity (page 40)

Assume the center of the target is the true score…



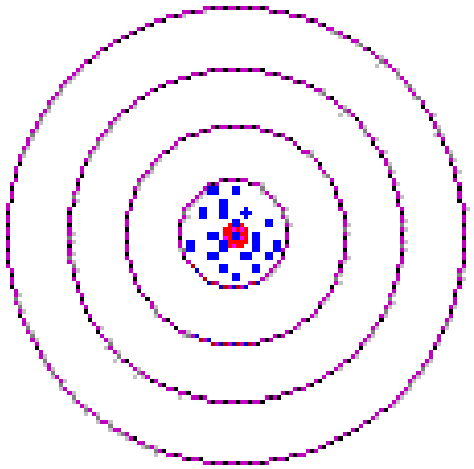
Reliable Not Valid

Consistent, but wrong



Neither Reliable Nor Valid

Inconsistent & wrong



Both Reliable And Valid

Consistent & correct

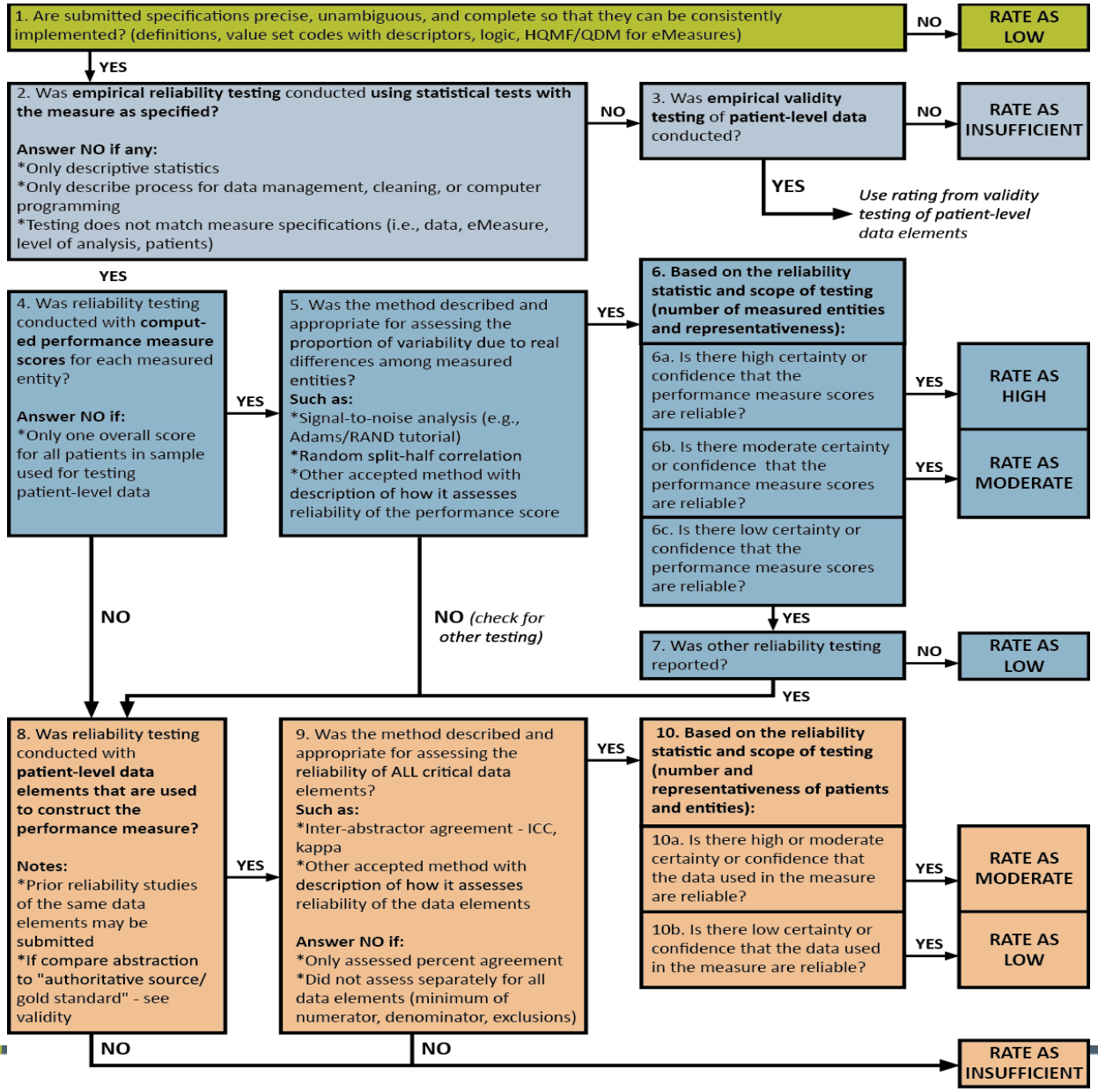
Evaluating Scientific Acceptability–Key Points (page 41)

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 42)

- 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

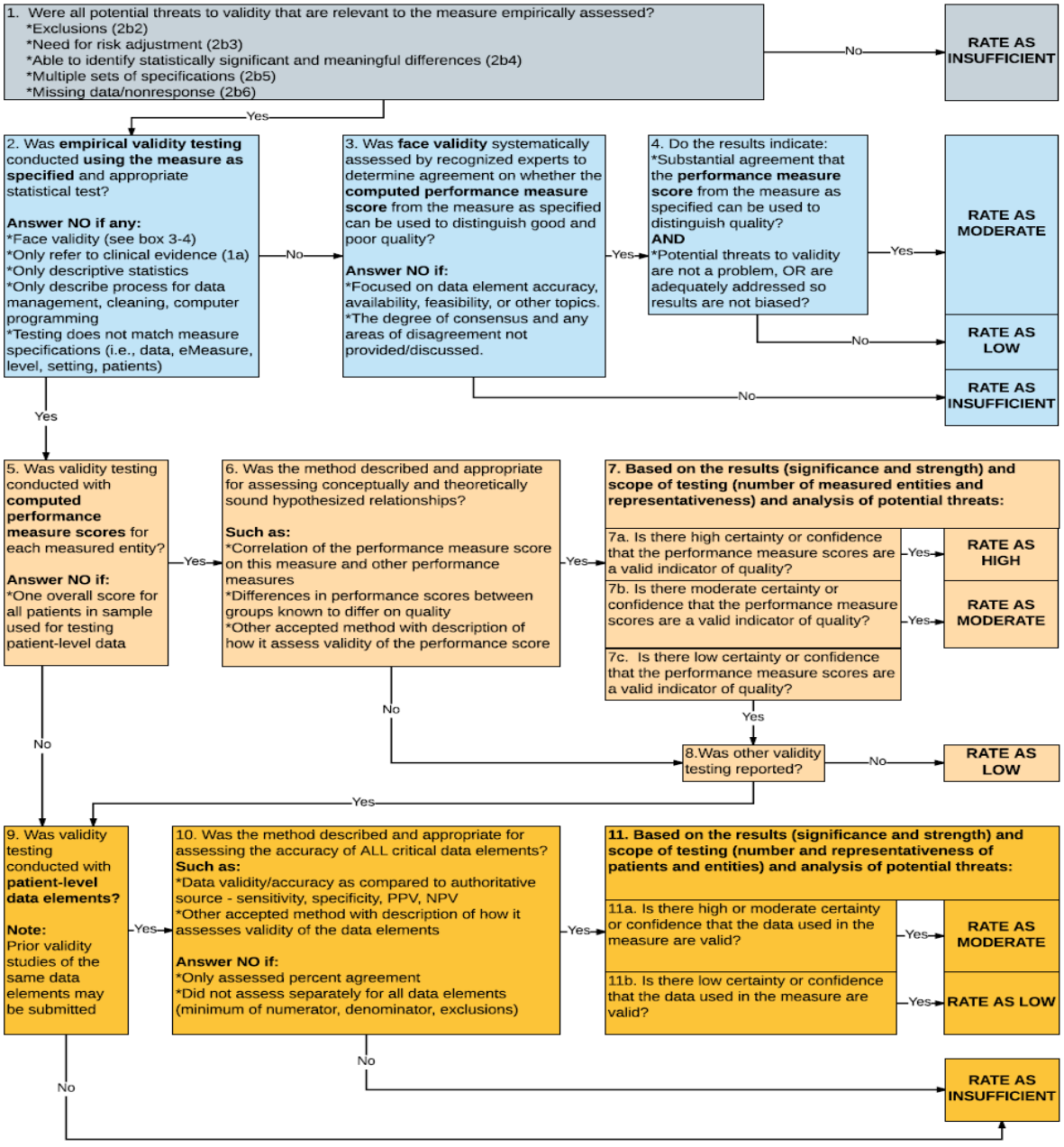
Rating Reliability: Algorithm #2 – page 43



Validity testing (pages 44-49)—Key points page 47

- **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.

Rating Validity: Algorithm #3 – page 48



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 49)—Key Points (page 50)

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 50)—Key Points (page 51)

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) Now 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

New measures	Maintenance measures
Feasibility	
• Measure feasible, including eMeasure feasibility assessment	NO DIFFERENCE: Implementation issues may be more prominent
Usability and Use	
• Use: used in accountability applications and public reporting • Usability: impact and unintended consequences	**INCREASED EMPHASIS**: Much greater focus on measure use and usefulness, including both impact and unintended consequences

Criterion #5: Related or Competing Measures (pages 51-52)

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.

- 5a. The measure specifications are harmonized with related measures **OR** the differences in specifications are justified.
- 5b. The measure is superior to competing measures (e.g., is a more valid or efficient way to measure) **OR** multiple measures are justified.

Updated guidance for measures that use ICD-10 coding: Fall 2017 and 2018

- Gap can be based on literature and/or data based on ICD-9 or ICD-10 coding
- Submit updated ICD-10 reliability testing if available; if not, testing based on ICD-9 coding will suffice
- Submit updated validity testing
  - *Submit updated empirical validity testing on the ICD-10 specified measure, if available*
  - *OR face validity of the ICD-10 coding scheme plus face validity of the measure score as an indicator of quality*
  - *OR face validity of the ICD-10 coding scheme plus score-level empirical validity testing based on ICD-9 coding*
  - *OR face validity of the ICD-10 coding scheme plus data element level validity testing based on ICD-9 coding, with face validity of the measure score as an indicator of quality due at annual update*

Evaluation Process

- **Preliminary analysis (PA)**: To assist the Committee evaluation of each measure against the criteria, NQF staff and the Scientific Methods Panel (if applicable) will prepare a PA of the measure submission and offer preliminary ratings for each criteria.
  - *The PA will be used as a starting point for the Committee discussion and evaluation*
  - *Methods Panel will complete review of Scientific Acceptability criterion for complex measures*
- **Individual evaluation**: Each Committee member conducts an in-depth evaluation on all measures
  - *Each Committee member will be assigned a subset of measures for which they will serve as lead discussant in the evaluation meeting.*

Evaluation Process

- **Measure evaluation and recommendations at the in-person/web meeting**: the entire Committee will discuss and rate each measure against the evaluation criteria and make recommendations for endorsement.
- **Staff will prepare a draft report** detailing the Committee's discussion and recommendations
  - *This report will be released for a 30-day public and member comment period*
- **Post-comment call**: the Committee will re-convene for a post-comment call to discuss submitted comments
- **Final endorsement decision by the CSAC**
- **Appeals** (if any)

Questions???

SharePoint Overview

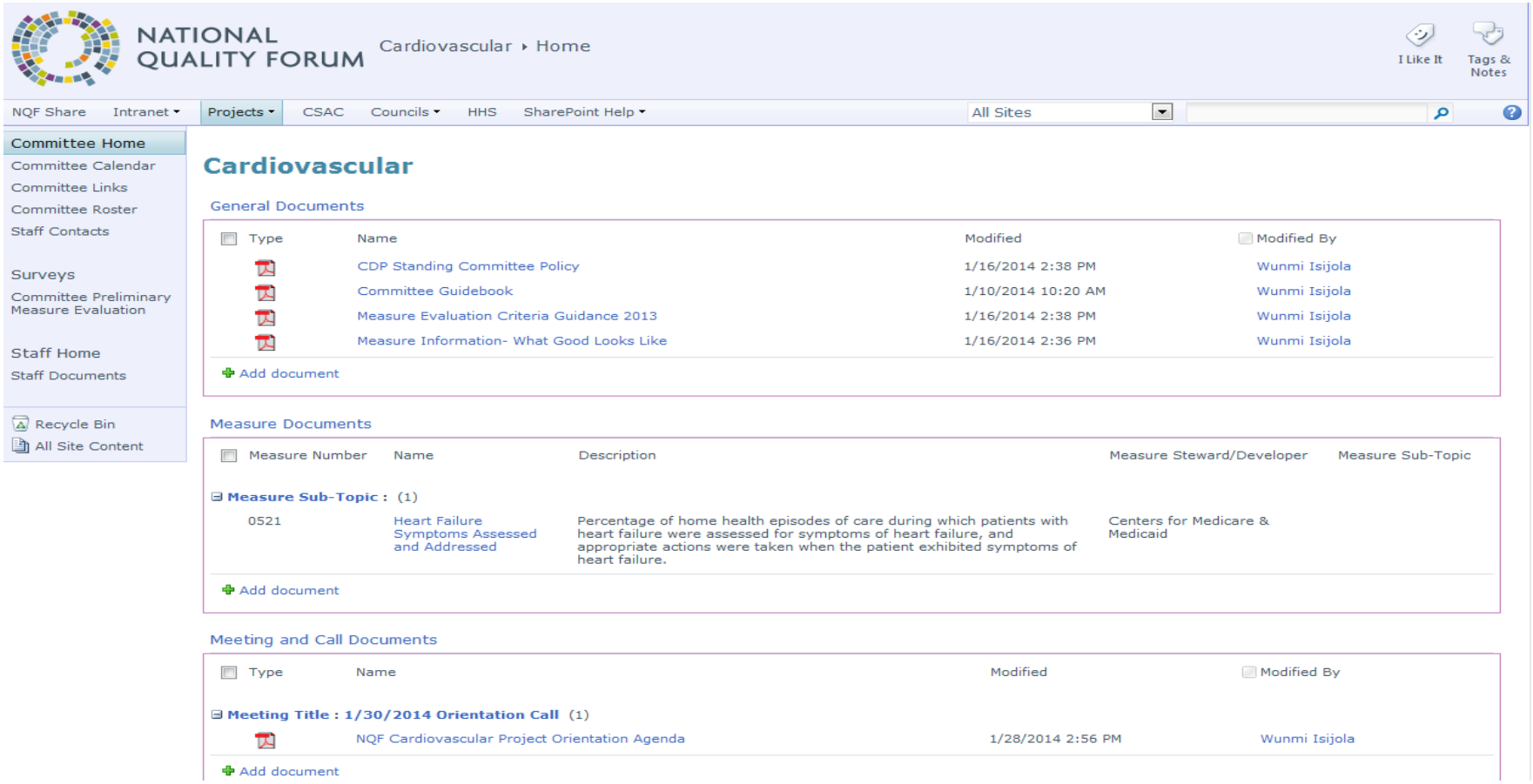
SharePoint Overview

http://share.qualityforum.org/Projects/Renal/SitePages/Home.aspx

- Accessing SharePoint
- Standing Committee Policy
- Standing Committee Guidebook
- Measure Document Sets
- Meeting and Call Documents
- Committee Roster and Biographies
- Calendar of Meetings

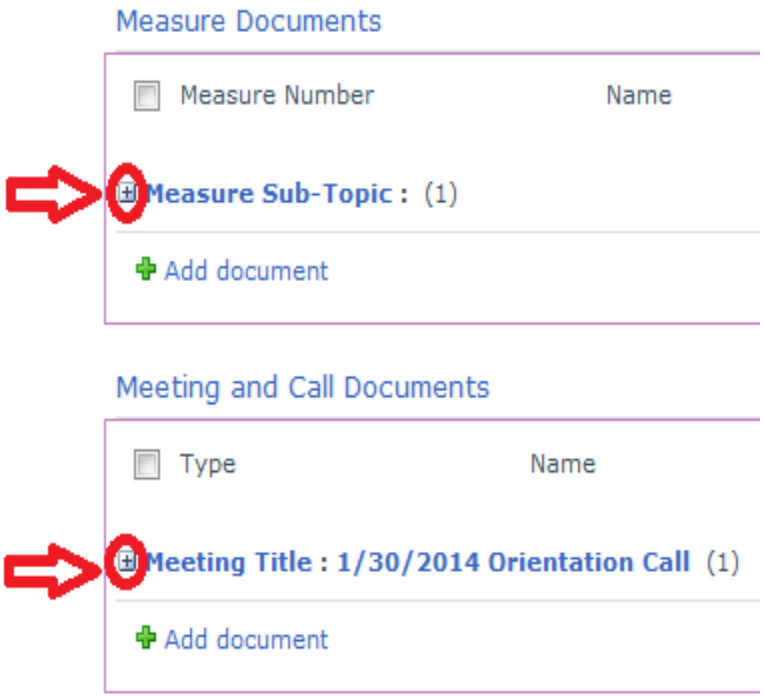
SharePoint Overview

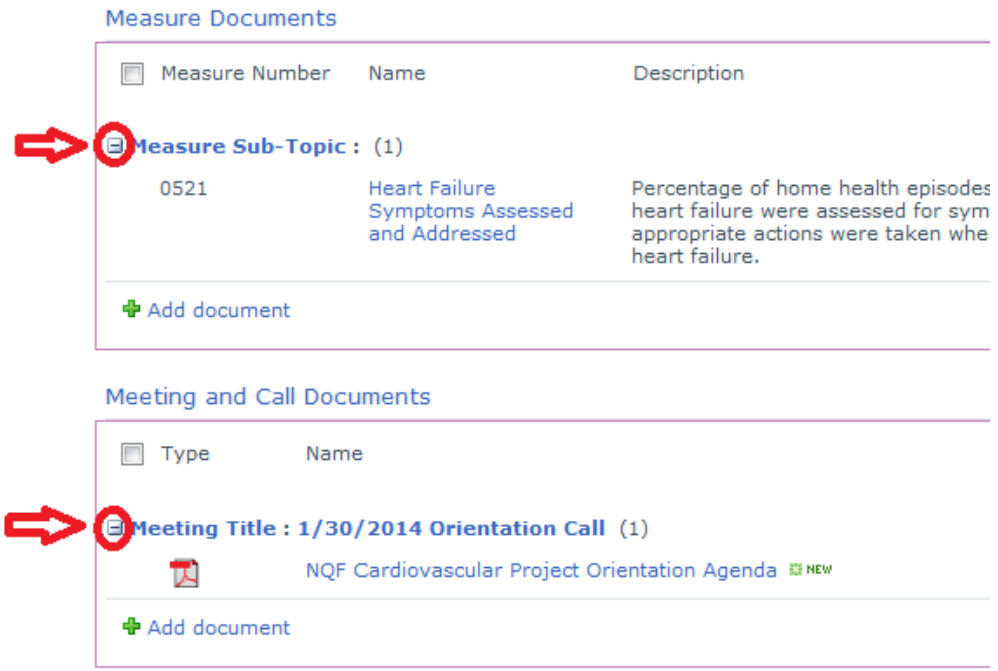
Screen shot of homepage



SharePoint Overview

Please keep in mind: + and – signs





Measure Worksheet and Measure Information

- Preliminary analysis, including eMeasure Technical Review if needed, and preliminary ratings
- Member and Public comments
- Information submitted by the developer
  - *Evidence and testing attachments*
  - *Spreadsheets*
  - *Additional documents*

Next Steps

Next Steps

Web Meetings

- Friday, June 15, 2018, 3:00 – 5:00 pm ET
- Monday, June 18, 2018, 3:00 – 5:00 pm ET
- Wednesday, June 19, 2018, 3:00 – 5:00 pm ET

Project Contact Information

- Email: Renal@qualityforum.org
- NQF Phone: 202-783-1300
- Project page: http://www.qualityforum.org/Project_Pages/Renal.aspx
- SharePoint site: http://share.qualityforum.org/Projects/Renal/SitePages/Home.aspx

Questions???


THANK YOU