National Consensus Standards for Cost and Resource Use: Phase II

Standing Committee Orientation

Evan Williamson
Ashlie Wilbon
Ann Phillips
Taroon Amin

NATIONAL QUALITY FORUM

Welcome & Introductions

NQF Project Staff

- Evan Williamson
  - Project Manager, Performance Measures
- Ann Phillips
  - Project Analyst, Performance Measures
- Taroon Amin
  - Senior Director, Performance Measures
- Ashlie Wilbon
  - Managing Director, Performance Measures

Standing Committee

- Brent Asplin, MD, MPH (Co-Chair)
- Lisa Latts, MD, MSPH, MBA, FACP (Co-Chair)
- Ariel Bayewitz, MPH*
- Lawrence Becker
- Mary Ann Clark, MHA
- Cheryl Damberg, PhD
- Jennifer Eames-Huff, MPH
- Nancy Garrett, PhD
- Andrea Gelzer, MD, MS, FACP
- Stanley Hochberg, MD
- Matthew McHugh, PhD, JD, MPH, RN
- Martin Marciniak, MPP, PhD
- James Naessens, ScD, MPH
- Eugene Nelson, DSc, MPH
- Jack Needleman, PhD
- Janis Orlowski, MD, MACP*
- Carolyn Pare
- John Ratliff, MD, FACS, FAANS*
- Andrew Ryan, PhD
- Joseph Stephansky, PhD
- Lina Walker, PhD
- William Weintraub, MD, FACC**
- Herbert Wong, PhD
- Dolores Yanagihara, MPH

*New member of the Committee

** Member of the Cardiovascular TEP

Cardiovascular Technical Expert Panel (TEP)

- William Weintraub, MD, FACC (Chair)*
- Sana Al-Khatib, MD, MHS
- Leslie Cho, MD
- Ted Gibbons, MD
- Judd Hollander, MD
- Thomas Kottke, MD, MSPH

* Member of the Cost and Resource Use Standing Committee

Agenda for the Call

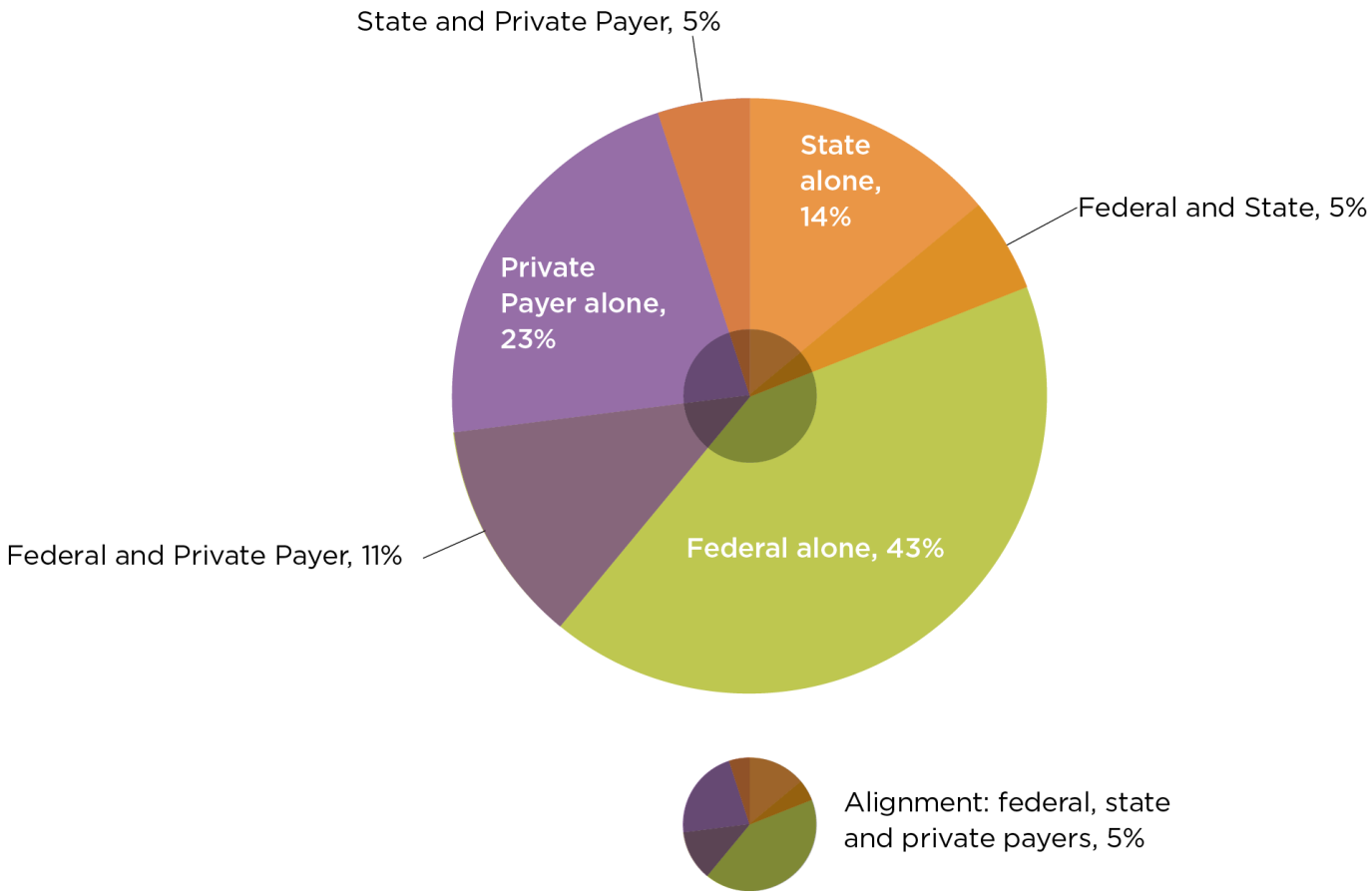
- NQF Overview
- Cost and Resource Use Portfolio of Measures
- Project Scope
- Role of the Committee and TEP
- SharePoint Tutorial
- Measure Evaluation Overview

NQF Mission


Board of Directors
Steering Committees
8 Membership Councils
Measures Application Partnership (MAP)
National Priorities Partnership (NPP)
CSAC, HITACH
Neutral Convener
Standards Setting Organization
1 Build Consensus
2 Endorse National Consensus Standards
3 Education and Outreach

Who Uses Endorsed Measures?

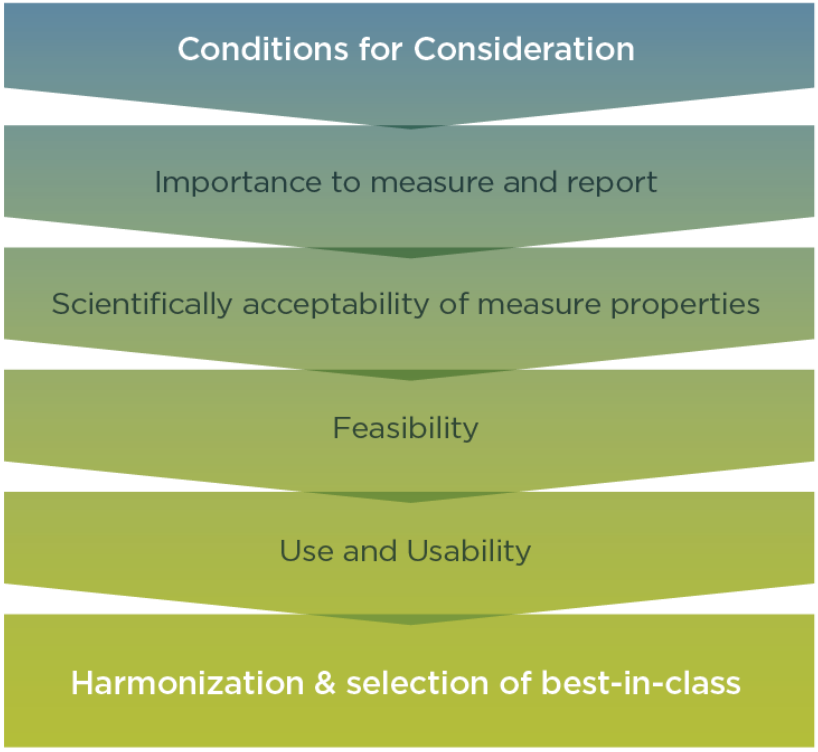
- Approximately 700 endorsed measures
- Various users
  - Federal
  - State
  - Community
  - Facility



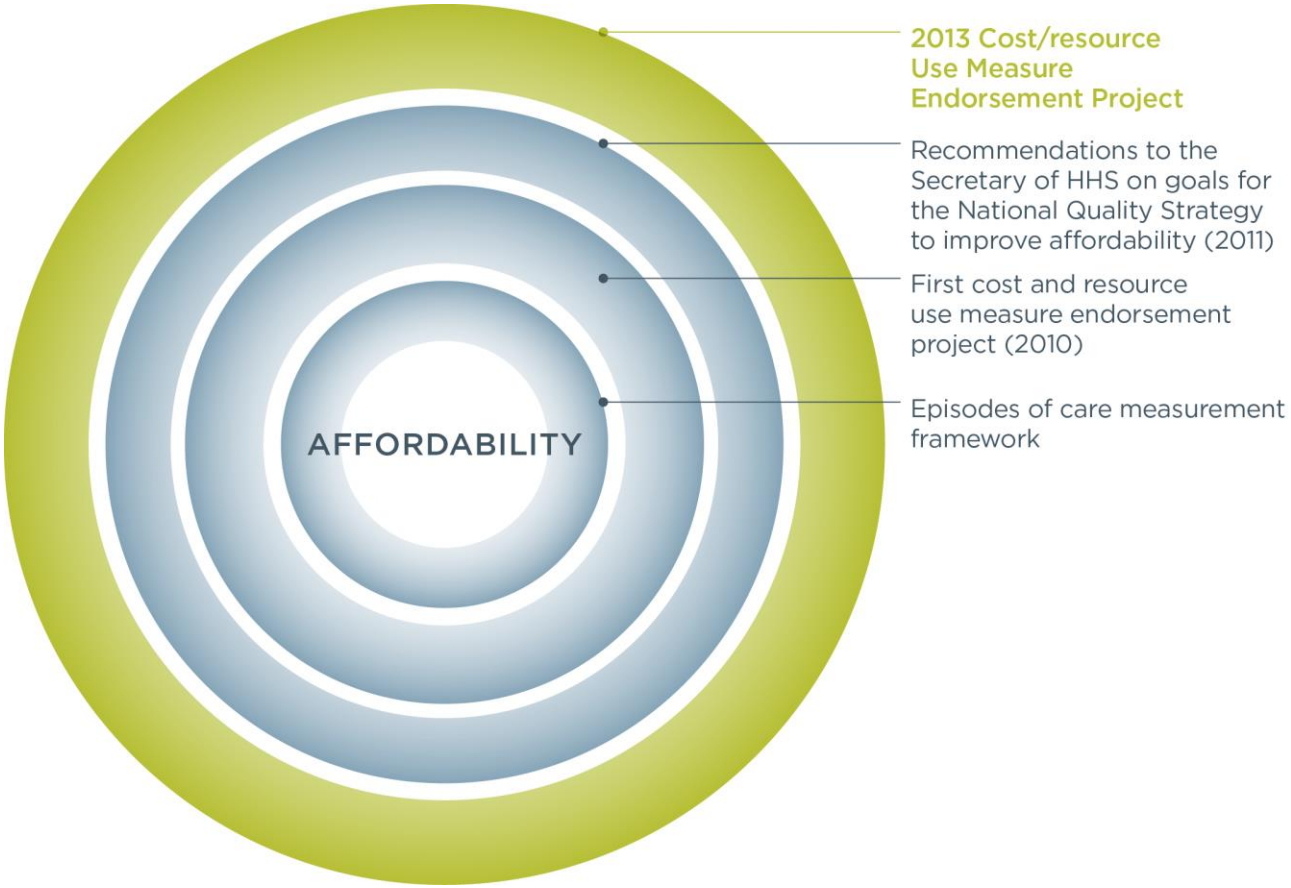
NQF Consensus Development Process (CDP)

8 Steps for Measure Endorsement

NQF Measure Evaluation Criteria



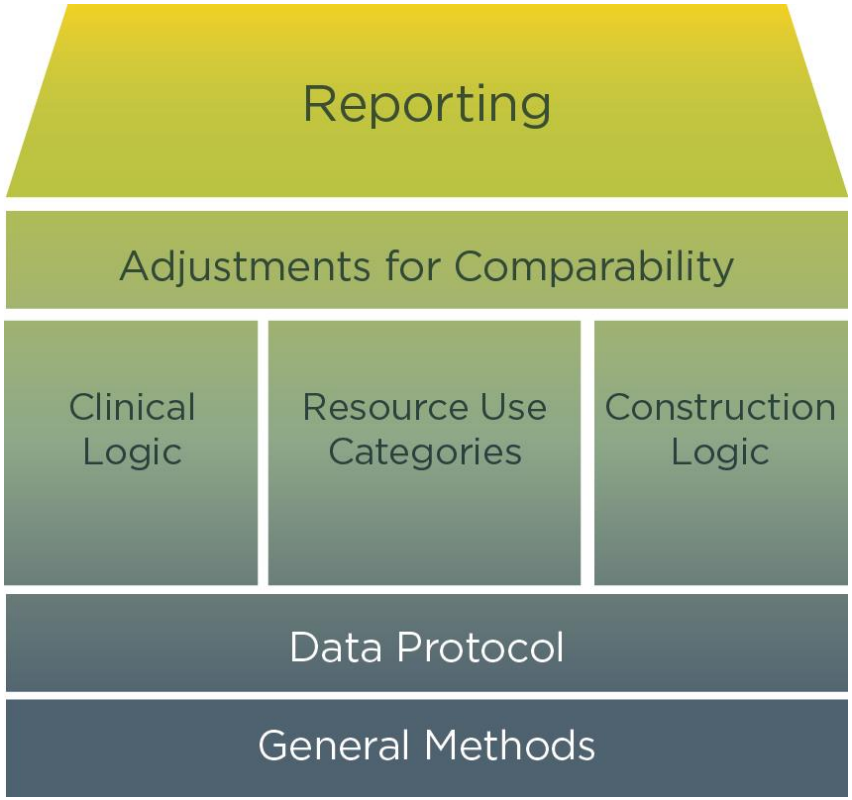
NQF and Cost and Resource Use:
Prior work



Defining Resource Use Measures

- Broadly applicable and comparable measures of health services counts (in terms of units or dollars) that are applied to a population or event (may include diagnoses, procedures, or encounters).
  - A resource use measure counts the frequency of defined health system resources; some further apply a dollar amount (e.g., allowable charges, paid amounts, or standardized prices) to each unit of resource.

Building Resource Use Measures



Resource Use: A Building Block

Endorsed Cost & Resource Use Measures

- Endorsed January 30, 2012:
  - 1598: Total Resource Use Population-based PMPM Index (HealthPartners)
  - 1604: Total Cost of Care Population-Based PMPM Index (HealthPartners)
  - 1558: Relative Resource Use for People with Cardiovascular Conditions (NCQA)*
  - 1557: Relative Resource Use for People with Diabetes (NCQA)
- Endorsed March 30, 2012:
  - 1560: Relative resource use for people with asthma (*NCQA*)**
  - 1561: Relative resource use for people with COPD (*NCQA*)**
  - 1609: ETG-based hip/knee replacement cost-of-care *(Ingenix)*
  - 1611: ETG-based pneumonia cost-of-care (*Ingenix*)**
- Endorsed December 6, 2013:
  - 2158: Medicare Spending per Beneficiary (MSPB) *(CMS)*

  - *Up for Maintenance in Phase 2
  - **Up for Maintenance in Phase 3

Cost & Resource Use Project

Phase 1: Total cost per capita and episode-based measures

- 2 measure submissions
  - 2158: Medicare Spending per Beneficiary (MSPB) – Endorsed December 2013
  - 2165: Standardized-Price Total Per Capita Per Beneficiary (FFS)-Not Endorsed

Phase 2: Cardiovascular Condition-Specific Measures

- 3 measure submissions
  - 1558: Relative Resource Use for People with Cardiovascular Conditions *(NCQA)**
  - 2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI) *(CMS/Yale)*
  - 2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for heart failure (HF) *(CMS/Yale)*

Phase 3: Pulmonary Condition-Specific Measures

- Measure Submission Deadline – April 18, 2014
  - 1560: Relative resource use for people with asthma (*NCQA*)*
  - 1561: Relative resource use for people with COPD (*NCQA*)*
  - 1611: ETG-based pneumonia cost-of-care (*Ingenix*)*
  - Pneumonia Measure Submission *(CMS)*

*Maintenance Measures

Role of the Steering Committee

- Act as a proxy for the NQF multi-stakeholder membership for a specific project
- Work with NQF staff to achieve the goals of the project
- Evaluate candidate measures against the measure evaluation criteria
- Respond to comments submitted during the review period
- Respond to any directions from the CSAC

Roles of the Steering Committee, cont.

- All Members review ALL measures
- Evaluate measures against each criterion
  - Indicate the extent to which each criterion is met and rationale for the rating
- Make recommendations to the NQF membership for endorsement

Role of the Steering Committee Co-Chairs

- Facilitate Steering Committee (SC) meetings
- Represent the SC at CSAC meetings
- Keep SC on track to meet goals of the project without hindering critical discussion/input
- Assist NQF in anticipating questions and identifying additional information that may be useful to the SC
- Work with NQF staff to achieve the goals of the project
- Participate as a SC member

Role of the Technical Expert Panel

- Review ALL submitted measures
- Provide input to the committee on the clinical specifications of the measures
- Answer questions from the committee regarding appropriateness of the clinical specifications

Role of NQF Staff

- **NQF project staff works with SC to achieve the goals of the project and ensure adherence to the consensus development process:**
  - Organize and staff SC meetings and conference calls
  - Guide the SC through the steps of the CDP and advise on NQF policy and procedures
  - Review measure submissions and prepare materials for conference review
  - Draft and edit reports for SC review
  - Ensure communication among all project participants (including SC and measure developers)
  - Facilitate necessary communication and collaboration between different NQF projects

Role of NQF Staff, cont.

- Respond to NQF member or public queries about the project
- Maintain documentation of project activities
- Post project information to NQF website
- Work with measure developers to provide necessary information and communication for the SC to fairly and adequately evaluate measures for endorsement
- NQF project staff works with communications department to publish final report

Activities and Timeline: Phase 2

Process Step	Timeline
Measure submission deadline	12/9/13
SC member orientation	1/13/14
SC and TEP Receive Measures	1/15/14
TEP member evaluation and review	1/15/14 – 1/31/14
TEP member submit evaluations on SharePoint	Due by 1/31/14
SC member preliminary review and evaluation	1/15/14 – 2/24/14
SC members submit evaluations online	Due by 2/24/14
SC in-person meeting	3/4/14 – 3/5/14
Draft report posted for NQF Member and Public Review and Comment	4/21/14 – 5/21/14
SC call to review and respond to comments	6/4/14 from 11am - 1pm ET
Draft report posted for NQF Member vote	6/18/14 – 7/2/14
CSAC review and approval	7/3/14 – 7/23/14
Endorsement by the Board	7/24/14 – 8/4/14
Appeals	8/5/14 – 9/4/14

NQF SharePoint Site

- http://share.qualityforum.org/Projects/costRU/SitePages/Home.aspx
- The SharePoint site will be the primary method of document sharing and collaboration for NQF Staff, Standing Committee, and Technical Expert Panel
- SharePoint Categories
  - Documents
  - Calendar
  - Committee and Staff Contacts
  - Evaluation Survey

NQF SharePoint Site

- Documents
  - General Documents
    - Standing Committee Guidebook, Roster and Bios, Measure Evaluation Criteria
  - Measure Documents
    - Measure Document Sets for each submitted Measure
  - Meeting and Call Documents
    - Agenda, Call information, and meeting materials for each conference call and the in-person meeting

Standing Committee Guidebook

- Newly developed for 2014
- Provides in-depth information on everything you will need to perform your role as a member of the committee
- Topics covered include -
  - Background on NQF
  - Evolving Performance Measurement Landscape
  - ABCs of Measurement
  - NQF Endorsement of Consensus Standards
  - Measure Evaluation Process
  - Measure Evaluation Criteria

Measure Evaluation Overview

Four Major Endorsement Criteria Hierarchy and Rationale

- Describe desirable characteristics of quality performance measures for endorsement
  - **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***)
  - **Scientific acceptability of measure properties** : Goal is to make valid conclusions about resource use; if not reliable and valid, there is risk of improper interpretation (***must-pass***)
  - **Feasible**: Goal is to, ideally, cause as little burden as possible; if not feasible, consider alternative approaches
  - **Usable**: Goal is to use for decisions related to accountability and improvement; if not useful, probably do not care if feasible
- If suitable for endorsement, evaluate measure harmonization and best-in-class

Criterion #1: Importance to Measure & Report (pg. 2)

Extent to which the measure focus is important to making significant contributions toward understanding healthcare costs for a high-impact aspect of healthcare where there is variation

1a. High Priority
The measure focus addresses: a specific national health Goal/Priority identified by DHHS or the National Priorities Partnership convened by NQF **OR** a demonstrated high-impact aspect of healthcare

1b. Opportunity for Improvement
Demonstration of resource use or cost problems, opportunity for improvement, or variations in care delivery

1c. Measure Intent
The intent of the resource use measure and the measure construct are clearly described **AND** the resource use service categories are consistent with the intent of the measure.

Criterion # 2: Reliability and Validity – Scientific Acceptability of Measure Properties (pg. 3)

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the cost or resources used to deliver care

2a. **Reliability (must-pass)**
- 2a1. Precise specifications including exclusions (previously 2d)
- 2a2. Reliability testing—data elements or measure score

2b. **Validity (must-pass)**
- 2b1. Specifications consistent with evidence
- 2b2. Validity testing—data elements or measure score
- 2b3. Justification of exclusions—relates to evidence
- 2b4. Risk adjustment
- 2b5. Identification of differences in performance
- 2b6. Comparability of data sources/methods

2c. Stratification for disparities

Measure Testing

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.

--Measure Testing Guidance Report

Evaluation of Testing

- Was an appropriate method used?
  - Consider level (data or score), data source, type of measure, topic, potential sources of error, conceptual relationships, feasibility
- Was the scope of testing adequate?
  - If sample, consider number of entities, number of patients, representativeness
- Were the results within acceptable norms?

Reliability Testing

- Reliability of ***data elements*** refers to repeatability and reproducibility of the data elements for the same population in the same time period.

- Examples
  - Inter-rater or intra-rater reliability (coding, record abstraction)
  - Internal consistency reliability for multi-item instruments

Reliability Testing

- 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)

Criterion #3: Feasibility (pg. 5)

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 (pg. 6)

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.

4a: Accountability: 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

4b: Improvement: Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated

4c: 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).

4d. Transparency: Data and result detail are maintained such that the resource use measure, including the clinical and construction logic for a defined unit of measurement can be deconstructed to facilitate transparency and understanding.

5. Comparison to Related or Competing Measures (pg. 7)

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.

Technical Expert Panel Review

Clinical Specifications of Cost and Resource Use Measures

- TEP Members will evaluate the clinical specifications from the following submission elements
- **Clinical Logic**
  - S.8.1. Brief Description of Clinical Logic
  - S.8.2. Clinical Logic
  - S.8.3. Evidence to Support Clinical Logic
  - S.8.4. Measure Trigger and End mechanisms
- **Adjustments for Comparability – Inclusion/Exclusion Criteria**
  - S.9.1. Inclusions and Exclusion
- **Adjustments for Comparability – Risk Adjustment**
  - S.9.3. Risk adjustment

Clinical Logic Evaluation

- **Clinical Logic**
  - To what extent is the measure population clinically appropriate?
  - To what extent are the definitions used to identify the measure population clinically consistent with the intent of the measure?
- **Evidence to Support Clinical Logic**
  - To what extent does the submission adequately describe the evidence that supports the decisions/logic for grouping claims (i.e., identifying the measure population, exclusions) to measure the clinical condition for the episode?
- **Measure Trigger and End mechanisms of the Episode**
  - Given the condition being measured, and the intent of the measure, describe the alignment of the length of the episode (including what triggers the start and end) with the clinical course of this condition.

Adjustments for Comparability – Inclusion/Exclusion Criteria

- **Clinical Inclusions and Exclusions**
  - Describe the clinical relevancy of the exclusions to narrowing the target population for the episode, condition/clinical course or co-occurring conditions, and measure intent.
  - Do the exclusions represent a large number or proportion of patients?
  - To what extent is the rationale for clinical exclusions adequately described and clinically relevant?
  - To what extent are the relevant conditions represented in the codes listed in the submission for clinical inclusions and exclusions?

Adjustments for Comparability – Risk Adjustment

- **Risk adjustment**
  - To what extent are the covariates (factors) included in the risk adjustment model clinically relevant and consistent with the measure's intent?

Questions?

NATIONAL QUALITY FORUM

Wrap-Up

Next Steps – Technical Expert Panel

- Technical Expert Panel Evaluation Conference Calls:
  - Monday, February 3rd 2-4pm ET ***AND***
  - Wednesday, February 12th 11am-1pm ET
- Committee and TEP Measure Evaluation Q&A Call:
  - Wednesday, February 19th 11am-1pm ET
- Complete your preliminary evaluation via SharePoint by:
  - COB on Monday, January 31st

Next Steps- Standing Committee

- Measure Evaluation Q&A Calls:
  - Monday, February 10th 11am-1pm ET
  - Wednesday, February 19th 11am-1pm ET (TEP will also be present)
- Complete your preliminary evaluation surveys via SharePoint by:
  - COB on Monday, February 24th
- In person Committee meeting:
  - Tuesday, March 4th and Wednesday, March 5th in Washington, DC

Project Contact Info

- Evan: ewilliamson@qualityforum.org
- Ashlie: awilbon@qualityforum.org
- Ann: aphillips@qualityforum.org
- Taroon: tamin@qualityforum.org
- NQF Phone: 202-783-1300

- SharePoint site:

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

Questions?

NATIONAL QUALITY FORUM