

## **Intended Use Advisory Panel**

Preliminary Recommendations  
to Enhance the Consensus  
Development Process (CDP)



**NATIONAL  
QUALITY FORUM**

# Agenda

- Introduction and Background
- Panel Charge
- Key Themes
- Recommendations

# Panel Members

- Helen Darling, MA, National Business Group on Health, Chair, NQF Board of Directors
- Lee A. Fleisher, MD, University of Pennsylvania, Chair, NQF Consensus Standards Approval Committee (CSAC)
- Cristie Upshaw Travis, MSHA, Memphis Business Group on Health, Immediate Past Chair, NQF CSAC
- Elizabeth E. Drye, MD, SM, Yale School of Medicine
- Don Goldmann, MD, Institute for Healthcare Improvement
- Kate Goodrich, MD, MHS, Centers for Medicare & Medicaid Services
- Bruce L. Hall, MD, MBA, PhD, BJC Healthcare
- Mary Beth Landrum, PhD, Harvard Medical School
- Beth A. McGlynn, PhD, MPP, Kaiser Permanente
- Jonathan Perlin, MD, MSHA, PhD, FACP, Hospital Corporation of America
- Andrew Ryan, PhD, University of Michigan
- Lina Walker, PhD, AARP Public Policy Institute
- Rachel Werner, MD, PhD, University of Pennsylvania

# Introduction

- NQF strives to continuously improve its processes to reflect the changing needs of the American healthcare system
- Over the past several years, stakeholder groups have questioned whether endorsement should consider the specific intended or actual use(s) of a measure in its evaluation (e.g., in particular federal quality programs), and more broadly, whether the specific use of a measure should be included in the NQF measure endorsement criteria
- This effort by the NQF Intended Use Advisory Panel seeks to consider the merit of and the various approaches to considering a measure's specific intended or actual use(s) as part of the measure endorsement process

# Background – Consensus Taskforce

In 2012, NQF's Board of Directors (BoD) empaneled a Consensus Task Force (CTF) that recommended that this Advisory Panel consider two potential new directions:

- **Endorsement of measures for a specific intended or actual use(s).** Consider endorsement for specific purposes (e.g., internal quality improvement, public reporting, payment), with consideration that measures may not be suitable for all potential uses
- **Levels of endorsement by measure grading** (independent of use): Levels of endorsement would allow NQF to apply the same criteria to all measures, but to grade endorsed measures differently

# Intended Use Advisory Panel Charge

The Charge of this Panel was to:

- Discuss several critical topic areas, including identifying various use cases for NQF-endorsed measures, distinguishing among the use cases, and identifying the need, if any, for different measure attributes, depending on the specific intended or actual measure use(s)
- Examine if the NQF measure endorsement criteria requires updating
- Propose a path forward on whether, and if so, how, to incorporate the specific use of measures in the endorsement process

# The key themes that emerged from the Advisory Panel deliberations included:

- There is limited consensus of the Advisory Panel on whether there is a hierarchy among the various accountability applications
- There is a qualitative difference between measures used for quality improvement (QI) only and those used in other applications
- Field experience could be a useful addition in the evaluation of measures used in accountability applications

# Theme 1: Limited consensus on whether there is a hierarchy among the various accountability applications

There were three differing views:

- A hierarchy could be determined based on the potential risk of misclassifying providers in the program and the subsequent financial impact to providers
- A hierarchy of financial impact to providers would be difficult to determine because most accountability applications have somewhat of a financial impact on providers
- A hierarchy based on financial impact to providers may be misguided, and that any hierarchy is inherently dependent on the perspective of the user



## Theme 2: There is a qualitative difference between measures used for quality improvement (QI) only

- Measures for QI-only are used by providers to evaluate their performance over time or among a group of providers and are not typically used publicly for comparative purposes
- Measures for internal QI may require a different standard for evidence and testing
- Different criteria for QI-only measures may help to create an environment where these organizations can more easily use and share measures.

## Theme 3: Field experience could be a useful addition in the evaluation of measures used in accountability applications

- Field experience data could include information on:
  - Gap identified from the measure as specified, or
  - An assessment of how well the measure identifies variation and meaningful differences of performance across providers.
- No consensus on the definition of what field experience would include
- Noted that it is important to create a mechanism to monitor how a measure is performing once it is implemented and to incentivize measure developers to provide these data for review.

## The recommendations that emerged from the Advisory Panel deliberations included:

1. Move forward with grading measures based on the NQF endorsement criteria
2. Incorporate “field experience” in the measure endorsement process
3. Consider the rapidly evolving uses of measures and incentivize grade advancement through further use and testing

## The recommendations that emerged from the Advisory Panel deliberations included: (cont.)

4. Consider different standards for QI measures
5. Encourage the Measures Applications Partnership (MAP) to consider how measure grades can be used in the selection of measures for programs
6. Pursue future work to consider the interaction between program attributes and individual measure attributes

# Recommendation 1: Move forward with grading measures based on the NQF endorsement criteria

- **AAA: Highest grade measures:**
  - Highest rating on all must pass endorsement criteria
  - High or moderate rating on feasibility, and usability and use.
  - Demonstrated field experience for at least one year
- **AA: High grade measures:**
  - High or moderate ratings for all NQF criteria
  - In contrast to AAA measures, there may be limited field experience with performance data.
- **A: Moderate grade measures:**
  - Moderate ratings for all NQF criteria
  - Limited field experience and performance data.

## Recommendation 2: Incorporate “field experience” in the measure endorsement process

- The Panel struggled to precisely define “field experience”
  - Goal: Ability to review performance data on the performance measure prior to widespread implementation
- Analysis could include:
  - Assessments of the gap in measure performance
  - An assessment of how well the measure identifies variation and meaningful differences of performance across providers
  - An assessment of the benefits and risks associated with the measure by end-users prior to widespread implementation
- The requirement for field experience would likely evolve with the maturity of the measure

## Recommendation 3: Consider the rapidly evolving uses of measures and incentivize grade advancement through further use and testing

- Reaffirmed the importance of identifying positive and negative unintended consequences when a measure is in use
  - Recognized the difficulty of accessing those data by measure developers
- Recommendation that the measure maintenance process be utilized as a collaborative opportunity to raise the measure's grade through further use and testing
  - NQF Measure Maintenance could create a positive incentive for measure developers to perform additional analyses or recommended updates on measures to move up the grading scale

## Recommendation 4: Consider different standards for QI measures

- Encouraged the development of a NQF process that would allow QI-only measures to be evaluated differently than other measures, and, potentially, approved
- While QI-measures should demonstrate opportunity for improvement or variation across providers, the review process could potentially lower requirements for the must-pass requirements of evidence, reliability, and validity, compared to the current endorsement process
- Since this would require a new process for NQF, **the Panel specifically seeks comment on the usefulness of an approval process for QI measures**



## Recommendation 5: Encourage the Measures Applications Partnership (MAP) to consider how measure grades can be used in the selection of measures for programs

- The Advisory Panel encourages the Measures Applications Partnership (MAP) to consider how the grades can be used when selecting individual measures for programs
- Example:
  - In an effort to align program and measure attributes, the MAP may determine that an individual program requires AAA measures, where as another program may only need AA measures.

## Recommendation 6: Pursue future work to consider the interaction between program attributes and individual measure attributes

- The Panel identified a preliminary set of measure attributes that can be considered, including:
  - Inclusion and exclusion criteria
  - Potential for misclassification based on reliability and validity testing results
  - The precision of the risk adjustment models.
- Further, a set of program attributes may include:
  - Methods used to define performance categories
  - Whether or not statistical tests are used to distinguish between performance categories and the approach to those tests
  - The nature of the financial incentive

# Request for Comment

**Comment is open through September 30, 6:00pm**

- The Advisory Panel seeks public comment on the key themes and the recommendations that emerged from its deliberations of the charge outlined by the NQF Board.
- The Panel will consider these comments during an upcoming comment call on October 20, 2015.
- Following this call, the recommendations will be finalized and presented to the Consensus Standards Approval Committee (CSAC) and the NQF Board.

# Timeline

- *Orientation Call: June 8 & June 10*
- *Call #2: Discussion of Possible Use – July 10*
- *Call #3: Possible Uses & NQF Criteria – July 29*
- **Public Comment Period: August 31- September 30, 2015**
- All-Member Town Hall Webinar: September 21, 1:00-2:30pm
- Post-Comment Call: October 20, 12:30-2:30pm ET
- CSAC: November 18-19 2015
- Executive Committee: December 8, 2016

# Project Contact Information

- Helen Burstin, Chief Scientific Officer
  - Taroon Amin, Senior Director-Consultant
  - Suzanne Theberge, Senior Project Manager
  - Poonam Bal, Project Manager
  - Kaitlynn Robinson-Ector, Project Analyst
- 
- Email: [intendeduse@qualityforum.org](mailto:intendeduse@qualityforum.org)
  - Project Page:  
[http://www.qualityforum.org/Intended Use.aspx](http://www.qualityforum.org/Intended_Use.aspx)