

REPORT FROM THE  
INTENDED USE ADVISORY PANEL

---

# Recommendations to Enhance the Consensus Development Process (CDP)

A FINAL REPORT  
FEBRUARY 11, 2016



NATIONAL  
QUALITY FORUM

# CONTENTS

EXECUTIVE SUMMARY	2
INTRODUCTION	4
BACKGROUND	4
INTENDED USE ADVISORY PANEL	5
PANEL RECOMMENDATIONS	5
APPENDIX A: Panel Members and NQF Staff	9

## EXECUTIVE SUMMARY

The National Quality Forum (NQF) strives to continuously improve its processes to reflect the changing needs of patients and the American healthcare system in which they get their care. Since its inception, one of the central principles of NQF's endorsement process has been that endorsed measures, having met rigorous criteria and approval from diverse stakeholders, should be suitable for use in both quality improvement and accountability (i.e., payment or public reporting) applications. However, over the past several years, some stakeholder groups have questioned whether NQF's endorsement process should consider the specific intended or actual use(s) of a measure.

This issue was considered by NQF's Consensus Task Force, which reviewed the endorsement process and recommended several specific enhancements to that process. Acting on the recommendations of the Consensus Task Force, NQF convened an Intended Use Advisory Panel in March 2015 to examine (1) whether and how the specific use of measures should be incorporated into the endorsement process; and (2) whether endorsement decisions should change from a simple yes/no to an incremental scale that would allow NQF to apply the same criteria to all measures, but to grade endorsed measures differently based on how well the measures met the criteria and, potentially, how well the measures performed when implemented. The Advisory Panel sought broad input from the public on its recommendations to enhance the consensus process. A revised version of this report reflecting this input was shared with the Consensus Standards Approval Committee (CSAC) and NQF Board for their consideration. Both the CSAC and the Board approved the revised recommendations.

A number of themes emerged from the Intended Use Advisory Panel's discussions. All members of the Panel agreed that measures that are

endorsed by NQF should produce reliable and valid information, regardless of the intended use of the measure. The Advisory Panel also agreed that providing more transparency to the public about the degree of a measure's reliability and validity would be very positive. Consequently, they recommend introducing an "NQF+" designation. To merit this new designation, a measure would have to exceed the current endorsement criteria in key areas and would have to be well-vetted in real-world settings by those being measured and other users.

There was general agreement that a qualitative difference between measures used for quality improvement (QI) only and those used in other applications exists and that NQF should focus on endorsement of measures intended for accountability applications. However, there was limited agreement among Panel members on whether measurement needs differ among the various accountability applications based on the risk of misclassification and the subsequent financial impact on providers. Consequently, the panel did not agree that the "NQF+" designation should correspond to a single accountability application, e.g., some on the panel argued that

the higher designation could apply to both pay for performance and public reporting applications. In sum, the Advisory Panel's recommendation is to introduce an "NQF+" designation, thereby increasing transparency, but to *not* associate the designation with a particular application.

The Advisory Panel encourages the Measure Applications Partnership (MAP) to consider whether the grades could be used in the selection of measures for various programs. Along these lines, the Panel recommended further examination of key measure and program attributes and their interaction to help inform both MAP and program implementers.

The Intended Use Advisory Panel made the following recommendations:

- Endorsement should not try to distinguish between the measures used in pay-for-performance, public reporting, and other types of accountability applications.
- Create an "NQF+" designation for endorsed measures that exceed the criteria for endorsement in key areas, and include a requirement for vetting by those being measured.

- NQF endorsement should focus on endorsement of measures intended for accountability applications, such as public reporting and payment applications.
- Encourage the Measure Applications Partnership (MAP) to consider how the NQF+ designation can be used in the selection of measures for programs.
- Pursue future work to consider the interaction between program attributes and individual measure attributes.

Following the publication of this report, NQF staff will begin working on implementation of these recommendations. Staff will integrate the changes into the Consensus Development Process (CDP) and will provide educational sessions and materials for NQF members, measure developers, and the public.

## INTRODUCTION

Over the past several years, various stakeholder groups have questioned whether NQF should incorporate the *specific* intended or actual use(s) of a measure (i.e., in particular federal quality programs) as part of the endorsement process, possibly via expansion of the NQF measure endorsement criteria. Doing so would move beyond the current Usability and Use criterion, which requires evaluation of the 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.

An effort by the NQF Intended Use Advisory Panel examined 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 Task Force

In 2012, NQF's Board of Directors empaneled a Consensus Task Force (CTF) that included members of the NQF Board, the Consensus Standards Approval Committee (CSAC), and representatives from the NQF membership. During its two-year tenure, the CTF reviewed the NQF endorsement process and recommended several enhancements. As part of its final set of recommendations, in the fall of 2014, the CTF advised the NQF Board to convene an Advisory Panel to consider transitioning from a binary endorsement decision (endorsed/not endorsed) to a more nuanced recommendation of endorsement. The CTF recommended that the Intended Use Advisory Panel consider two potential options.

1. **Endorsement of measures for a specific intended or actual use(s):** Endorsement for specific purposes (e.g., internal quality improvement, public reporting, payment) would assume that all measures may not be suitable for all potential uses. Currently, NQF endorses measures for quality improvement

and accountability applications. Accountability applications generally have been defined as the use of performance measures for comparative purposes, rather than solely for internal quality improvement. Endorsement for intended use would potentially allow NQF to hold measures used for different purposes to different standards and would recognize that different stakeholders may have different priorities.

2. **Distinguish levels or grades of endorsement (independent of use):** Levels of endorsement would allow NQF to apply the same criteria to all measures, but to grade endorsed measures differently. Endorsement would change from a simple yes/no to an incremental scale; measures could move up or down the scale, with periodic evaluations to maintain endorsement based on additional testing or experience. Measures could be graded based on how well they meet NQF endorsement criteria and how well they perform in actual use.

## INTENDED USE ADVISORY PANEL

Acting on the recommendations of the CTF in December 2014, the NQF Board approved the convening of an Intended Use Advisory Panel. Members of this Panel were selected to represent expertise in all aspects of quality measurement, including measure developers, program implementers, providers, and others who are measured by or use NQF-endorsed performance measures.

The Intended Use Advisory Panel was charged with:

- discussing 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);

- determining whether the NQF measure endorsement criteria require updating; and
- proposing a path forward on whether, and if so, how, to incorporate the specific use of measures in the endorsement process.

Accordingly, this Advisory Panel did not design new processes nor develop new policies for NQF. Instead, its contributions should be viewed as foundational and strategic recommendations that will initiate and facilitate continued dialogue with NQF members and other stakeholders on this topic via various public commenting opportunities. The Advisory Panel met virtually between June and October 2015. It considered public and member comments received prior to finalizing its recommendations and presenting them to the CSAC and NQF Board.

## PANEL RECOMMENDATIONS

The final five recommendations that emerged from the Advisory Panel deliberations are summarized below.

### RECOMMENDATION 1

Endorsement should not try to distinguish between the measures used in pay-for-performance, public reporting, and other types of accountability applications.

In considering whether NQF should endorse measures for specific use, the Advisory Panel examined at length the question of whether measurement needs differ among the various accountability applications (e.g., pay-for-performance, public reporting).

Several Panel members argued that there are differences based on the potential risk of misclassifying providers in particular programs

and therefore in the subsequent financial impact to providers. Some Panel members noted that measures used for pay-for-performance programs, particularly programs that are designed as penalty programs, should demonstrate the highest adherence to requirements for measure testing and evidence. Members agreed that there is a considerable impact to the health care system if measures used in payment programs

inappropriately misclassify providers and take away resources needed for patient care, as well as for performance improvement.

Other Panel members argued that most accountability applications have a financial impact on providers, regardless of whether they have direct penalties. They noted that pay-for-performance programs, but also public reporting, health plan network design, and other programs, have the potential to have a financial impact on providers. The use of performance measures can also have a financial impact on providers by informing consumer choice on the selection of providers and directing patients away from low-performing providers.

Others argued that identifying the measurement needs of these applications based on financial impact to providers might be misguided. They maintained that the risk of misclassification of providers has both a financial impact and a quality-of-care impact for patients, consumers, and purchasers. These panel members noted that the scientific rigor used to examine candidate measures should be the same regardless of the particular use, as many of the accountability applications drive purchasing and selection decisions by purchasers and consumers. They believe that NQF-endorsed measures should produce reliable and valid information, regardless of the intended use of the measure.

## RECOMMENDATION 2

Create an “NQF+” designation for endorsed measures that exceed the criteria for endorsement in key areas, and include a requirement for vetting by those being measured.

The Advisory Panel recommended introducing an “NQF+” designation for measures that have exceeded the criteria for endorsement in key areas. The following characteristics should be used to identify measures that achieve the “NQF+” designation:

- Meets evidence for measure focus without an exception;
- Measure is reliable as demonstrated by reliability testing of the measure score;
- Measure is valid as demonstrated by empirical validity testing of the measure score (i.e., not via face validity only); and
- Measure is well-vetted in real world settings by those being measured and other users.

The first three elements needed for the NQF+ designation signify that a measure meets a higher bar for several objectively evaluated elements of the NQF endorsement criteria. To ensure greater consistency in Standing Committee ratings, the

Panel recommended that NQF technical staff/consultants provide a preliminary analysis and rating on those must-pass criteria using detailed algorithms for rating (i.e., evidence, reliability, and validity). Increased transparency by using the “NQF+” designation will allow users to understand how well the measures meet these key criteria.

In addition to exceeding the criteria for endorsement in key areas, the Panel recommended including an evaluation of the measure experience in order to confer the “NQF+” designation. Specifically, this new experience element can help support the validity of the measure by demonstrating that a measure has been vetted in real world settings by those being measured, ensuring that the measure can be adopted and sustained in diverse contexts without undue burden or unintended negative consequences. This can be demonstrated when those being measured have been given:

- measure performance results and data,

- assistance with interpreting the measure results and data, and
- an opportunity to provide feedback on the measure performance and implementation by the community of entities being measured and other users.

The CSAC noted that this input should help avoid unintended consequences once measures are in use. Although NQF staff will identify those candidate measures that meet the “NQF+” designation, standing committees will make the final decision regarding endorsement recommendations, including conferring of the NQF+ designation.

The CSAC expressed concerns that the “NQF+” label may lead to confusion and that the name

should be reconsidered. Specifically, there were concerns about maintaining the creditability of regular endorsement. The naming convention of “NQF+” was chosen as a placeholder to signify that the measure has not only met NQF criteria for endorsement but exceeded them. NQF will be considering an alternative name that emphasizes the importance of both levels of endorsement. In this new endorsement designation, NQF will emphasize the value of transparency of the ratings on the criteria. Although few measures will likely achieve NQF+ designation to start, it is hoped that this change will demonstrate NQF’s commitment to the evolution of measures and drive toward increased reliability and validity testing. It will also help move NQF toward closing the feedback loop on measure use.

### RECOMMENDATION 3

#### NQF endorsement should focus on endorsement of measures intended for accountability applications, such as public reporting and payment applications.

The Panel recommended that the NQF endorsement process should support the use of measures for accountability applications and performance improvement. The Panel agreed that measures that are endorsed for accountability applications can be used for internal quality improvement (QI); however, NQF should make it clear that it does not endorse measures for 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 comparison. The Panel agreed that QI-only measures are exceedingly important and

that efforts to develop meaningful measures that facilitate both local improvement and shared learning should be accelerated. However, it did not support development of a new process by NQF to endorse or otherwise approve QI-only measures. The Panel acknowledged those public and member comments that supported use of the National Quality Measures Clearinghouse, developed by the Agency for Healthcare Research and Quality (AHRQ), as a resource for identifying measures suitable for QI only.



## RECOMMENDATION 4

---

Encourage the Measure Applications Partnership (MAP) to consider how the NQF+ designation can be used in the selection of measures for programs.

The Advisory Panel encouraged the Measure Applications Partnership (MAP) to consider how it could use the “NQF+” designation when selecting individual measures for specific programs. For example, in an effort to align program and measure attributes, MAP may determine that an

individual program requires “NQF+” measures. The Advisory Panel generally agreed that the MAP Coordinating Committee would be most appropriate to develop an approach for applying the “NQF+” designation in their future work.

## RECOMMENDATION 5

---

Pursue future work to consider the interaction between program attributes and individual measure attributes.

The Advisory Panel urged NQF to pursue future work to define key measure attributes and program attributes, examine their interaction, and give program implementers, purchasers, and health plans guidance on which measures may be better suited for implementation in specific programs based on program characteristics. The Panel identified a preliminary set of measure attributes that should be considered: (1) inclusion and exclusion criteria, (2) potential for misclassification based on reliability and validity testing results, and (3) the precision of the risk adjustment models. Further, they identified an initial set of program attributes to consider: (1) methods used to define performance categories (e.g., measure score thresholds), (2) whether or not statistical tests are used to distinguish between performance categories and the approach to those tests, and (3) the nature of the financial incentive (e.g., tied to performance or improvement, upside or downside risk or both, etc.).

This work should begin with identifying and categorizing measure and program attributes, and then move further to provide guidance on the interaction of these various elements. Such work would advance health care measurement science by identifying more precisely the likelihood that a measure will perform well in the context of a given program design. An examination of measure attributes and program attributes likely would reveal other key principles for the endorsement and application of measures. In turn, these key principles can help to inform the creation of a test environment in which measures and programs may be more precisely matched to drive health system performance improvement.

## APPENDIX A:

### Panel Members and NQF Staff

#### PANEL MEMBERS

---

**Helen Darling, MA**

National Business Group on Health  
Washington, DC

**Elizabeth E. Drye, MD, SM**

Yale School of Medicine, Center for Outcomes  
Research & Evaluation  
New Haven, CT

**Lee A. Fleisher, MD**

Perelman School of Medicine, University of  
Pennsylvania  
Philadelphia, PA

**Don Goldmann, MD**

Institute for Healthcare Improvement  
Cambridge, MA

**Kate Goodrich, MD, MHS**

Centers for Medicare & Medicaid Services  
Washington, DC

**Bruce L. Hall, MD, MBA, PhD**

BJC Healthcare  
Saint Louis, MO

**Mary Beth Landrum, PhD**

Harvard Medical School  
Boston, MA

**Beth A. McGlynn, PhD, MPP**

Kaiser Permanente Center for Effectiveness & Safety  
Research  
Pasadena, CA

**Jonathan Perlin, MD, MSHA, PhD, FACP**

Hospital Corporation of America  
Nashville, TN

**Andrew Ryan, PhD**

University of Michigan Center for Health Outcomes  
and Research  
Ann Arbor, MI

**Cristie Upshaw Travis, MSHA**

Memphis Business Group on Health  
Memphis, TN

**Lina Walker, PhD**

AARP Public Policy Institute  
Washington, DC

**Rachel Werner, MD, PhD**

University of Pennsylvania  
Philadelphia, PA

#### NQF STAFF

---

**Helen Burstin, MD, MPH**

Chief Scientific Officer

**Taroon Amin, PhD, MPH**

Consultant

**Suzanne Theberge, MPH**

Senior Project Manager

**Poonam Bal, MSHA**

Project Manager

**Kaitlynn Robinson-Ector, MPH**

Project Analyst

The Panel would like to acknowledge **Reva Winkler, MD, MPH**; **Karen Johnson, MS**; **Elisa Munthali, MPH**; and **Marcia Wilson, PhD** for the input and comments that they provided on the draft and final recommendations.

NATIONAL QUALITY FORUM  
1030 15TH STREET, NW, SUITE 800  
WASHINGTON, DC 20005  
[www.qualityforum.org](http://www.qualityforum.org)