

# NATIONAL QUALITY FORUM

## HEALTHCARE DISPARITIES AND CULTURAL COMPETENCY CONSENSUS STANDARDS

### IN-PERSON MEETING OF THE STEERING COMMITTEE

July 11-12, 2011

*Committee Members Present:* Dennis Andrulis, PhD, MPH (Co-Chair); Denice Cora-Bramble, MD, MBA (Co-Chair); Evelyn Calvillo,<sup>1</sup> DNSc, RN; Marshall Chin, MD, MPH; Luther Clark, MD; Lourdes Cuellar, MS, RPh; Colette Edwards, MD, MBA; Leonard Epstein, MSW; Dawn Fitzgerald,<sup>2</sup> MBA; Romana Hasnain-Wynia, PhD; Edward Havranek, MD; Elizabeth Jacobs, MD, MAPP; Jerry Johnson, MD; Francis Lu, MD; Mary Maryland, PhD, MSN; William McDade, MD, PhD; Ernest Moy, MD, MPH; Marcella Nunez-Smith, MD, MHS; Sean O'Brien, PhD; Norman Otsuka, MSc, MD; Grace Ting, MHA; Donna Washington, MD, MPH; Ellen Wu, MPH; Mara Youdelman, JD, LLM

*NQF Staff Present:* Helen Burstin, MD, MPH, Senior Vice President, Performance Measures; Heidi Bossley, MPH, Senior Vice President, Performance Measures; Nicole McElveen, MPH, Senior Project Manager; Robyn Y. Nishimi, PhD, NQF Consultant; Elisa Munthali, MPH, Senior Project Manager; Kristin Chandler, MPH, Project Analyst

*Others Present:* Joel S. Weissman, PhD, Massachusetts General Hospital; Joseph R. Betancourt, MD, MPH; Aswita Tan-McGrory, MPH

The full transcripts and audio recordings from the meeting can be found on the [project page](#).

## WELCOME AND INTRODUCTIONS

Ms. McElveen commenced the meeting and welcomed the Steering Committee. The Co-Chairs made welcoming comments and asked the Committee to disclose any conflicts of interest pertaining to the information under discussion. No disqualifying conflicts were reported; however, several members reported current/past involvement with grants and workgroups that may overlap with topics discussed during the meeting.<sup>3</sup>

Ms. McElveen provided a brief overview of the project and meeting objectives: 1) obtain recommendations for identifying disparities-sensitive measures; 2) obtain recommendations for identifying methodological considerations for measuring and reporting disparities; and 3) obtain guidance on how the National Quality Forum (NQF) should assess disparities and performance

---

<sup>1</sup> Evelyn Calvillo participated via phone.

<sup>2</sup> Dawn Fitzgerald participated via phone.

<sup>3</sup> Elizabeth Jacobs serves on Aetna's Racial Ethnic Disparities Advisory Task Force. Romana Hasnain-Wynia serves on the AHRQ Technical Advisory Panel for Disparities and Quality and the Quality Gap. Marcella Nunez-Smith is current principal investigator for a grant funded by the National Institutes of Health to develop a measure of healthcare discrimination. Mara Youdelman is chair of the Certification Commission of Healthcare Interpreters, which has developed competency standards for assessing interpreters in healthcare settings. Donna Washington conducts health disparities research on determinants of healthcare disparities within the VA Health System. Ed Havranek is principal investigator for a grant examining the effects of bias and discrimination on healthcare outcomes.

# NATIONAL QUALITY FORUM

measurement prospectively as new candidate standards are reviewed and/or existing standards undergo endorsement maintenance. Ms. McElveen also recapped the project timeline and provided an overview of NQF's historical work on disparities and cultural competency.

## **PRESENTATION: DRAFT COMMISSIONED PAPER ON HEALTHCARE DISPARITIES MEASUREMENT**

Drs. Weissman and Betancourt from The Disparities Solutions Center at Massachusetts General Hospital presented the draft commissioned paper addressing healthcare disparities measurement. The paper was specifically developed for this project to provide context, to identify the range of technical and policy issues pertinent to measuring healthcare quality disparities, and to help inform the Steering Committee's discussions.<sup>4</sup> The paper also was intended to inform the Committee's deliberations about how NQF could approach future efforts to endorse performance measures for disparities and cultural competency, including guidance that NQF can provide to measure developers. NQF and the authors had previously agreed that the paper would focus on racial, ethnic, and language proficiency disparities because of the quality of the evidence surrounding those specific types of disparities.

Drs. Weissman and Betancourt provided a general overview of the paper. The paper included five sections as outlined below:

- Background: Disparities and Quality Measurement
- Data Collection: Building the Foundation
- Disparities Measures and Indicators
- Methodological Approaches to Disparities Measurement
- Priorities and Recommendations for Quality Improvement and Public Reporting

## **DISCUSSION: HEALTHCARE DISPARITIES MEASUREMENT**

Discussion of the commissioned paper centered mainly on criteria for NQF endorsement and methodological approaches to disparities measurement and reporting (sections 3 and 4). The paper presented specific recommendations for each section, and the Committee was provided questions to consider during the review and discussion of the paper and its recommendations. Drs. Weissman and Betancourt provided a recap for each section prior to the Committee discussion and responded to queries from the Committee as they arose.

### **Disparities Measures and Indicators: What to Measure?**

#### ***Selection Criteria (Sections 3a and 5c)***

The paper recommended an expansion of the current NQF disparities-sensitive criteria (prevalence, impact of condition, impact of quality process, quality gap, ease and feasibility). It suggested a specific focus on prevalence and quality gap, noting the other criteria were more general and necessary for all measures, not just those focused on disparities. In addition, the

---

<sup>4</sup> The paper was revised based on the comments and suggestions from the Committee during this meeting. The revised draft is available for viewing on the NQF website.

# NATIONAL QUALITY FORUM

paper proposed focusing on four situations related to evidence so as to refine the criteria: 1) known disparities exist either currently or in the past for a specific (or similar) measure/aspect of care/outcome; 2) no data exist on disparities for a particular measure; 3) data exist and show no disparities; or 4) known disparities exist, but no quality measure exists. The paper also recommended the application of four additional criteria to provide greater specificity to a disparities-sensitive set: 1) care with a high degree of discretion; 2) communication-sensitive services; 3) lifestyle changes; and 4) outcomes rather than process measures.

The Committee generally agreed with the concept of reviewing the NQF portfolio of measures through the lens of these criteria to identify a subset of measures addressing disparities related to race and ethnicity. In addition, it also agreed with the report's recommendations for a primary focus on prevalence and quality gap and the application of "more subjective criteria." Suggestions were made for expanding the proposed criteria (mentioned earlier) to address impact on different levels (i.e., stakeholders, community, and end users), sample size and precision issues, and the context of community, social, and political issues. The Committee raised concerns about the criterion to focus on outcome rather than process measures, because outcome measures often present more issues related to risk adjustment, which could mask quality of care issues. Additionally, the Committee discussed the notion of an outcome measure serving as an indicator for provider accountability for a particular patient outcome. It was recognized that certain selection criteria may be difficult to measure and use as identifiers for disparities, such as impact and actionability. Lastly, the Committee noted that when selecting disparities-sensitive measures, an approach of "opt in" versus "opt out" might be most appropriate—that is, it might be easier to first assume all measures as disparities-sensitive and then use the selection criteria to remove a measure.

## ***Disparities-Specific Measures (Section 3c)***

The concept of disparities "sentinel" measures was introduced by the authors—that is, indicators for which research has shown disparities exist but no quality measures are available, and therefore a new measure would need to be developed to address a particular area. The draft reviewed by the Committee also made reference to sentinel measures as being those that might be in use yet serve as "sentinel" markers for quality disparities. The paper indicated that these measures could play a role in making healthcare organizations aware of disparities that may exist even though they may not be apparent when using standard quality indicators.

After much discussion clarifying what the authors intended with the use of "sentinel" (i.e., areas/indicators where no performance measures exist), the Committee agreed with and recommended that NQF adopt the overall approach of disparities "sentinel" measures. However, Committee members expressed concern with the overlapping meaning of "sentinel" because it is also used by the Joint Commission in a different context entirely. They suggested revising the word "sentinel" but maintaining the concept. The authors incorporated this recommendation into the draft paper that is posted on the NQF website.

## ***Measuring and Categorizing Disparities-Sensitive Measures (Section 3b)***

The paper proposed a categorization approach once disparities-sensitive measures are identified. Such an organizing schema would serve as a means to understand the range of NQF-endorsed<sup>®</sup>

# NATIONAL QUALITY FORUM

disparities measures and identify gaps in the portfolio. The categorization considers practitioner performance measures; consumer surveys that measure the patient experience; hospital, ambulatory care, home health, and nursing home measures; measures of ambulatory care sensitive conditions; measures of cultural competency; and patient-centered measures.

The Committee briefly discussed the various categories and agreed with the proposed categorization scheme. Committee members raised the issue of care coordination, questioning where it would fit in the organizational context. In addition, Committee members were concerned about the category of “root of the disparity,” as provider-based, patient-based, or systematic/health insurance. One Committee member cautioned about referencing the root cause of a disparity as being patient-based. The paper’s authors clarified that this specific designation was intended to indicate where interventions could be targeted, not necessarily to pinpoint a specific cause or assign blame for the aspect causing or contributing to the disparities.

## **Methodological Approaches to Disparities Measurement (Section 4)**

The Committee reviewed several methodological approaches and concepts for measuring disparities. The paper provided recommendations for each concept presented, allowing the Committee to discuss and develop overarching principles and specific criteria for disparities measurement and reporting.

### ***Reference Point (Section 4b)***

The reference point serves as the specific value against which a disparity is measured. The paper recommended that the reference point should always be the historically advantaged group, not the largest or best performing in an area/on a measure. The Committee generally agreed with this recommendation, noting that this approach would be consistent with that of the Agency for Healthcare Research and Quality (AHRQ) in its National Healthcare Quality and Disparities Reports. These reports maintain a fixed reference across all comparisons. In addition, the Committee discussed the regional implication of historically advantaged groups. Because the variances of this group may differ by geographic region, the Committee thought it could be important to account for these differences when comparisons are made.

### ***Absolute versus Relative Disparities and Favorable versus Adverse Measures (Section 4c)***

The absolute and relative changes in disparities can reveal different conclusions about whether gaps are actually closing and often can lead to different interpretations when making these comparisons. The paper recommended that both absolute and relative statistics should be calculated, and if they lead to conflicting conclusions, then both statistics should be presented, allowing readers to reach their own conclusion. The paper also notes that public reporting of disparities should calculate statistics using both favorable and adverse events; recommending that if the results are notably different, then both statistics should be reported—again, allowing readers to judge the importance by taking the context of the results into consideration.

The Committee did not favor one approach over another and generally agreed with the recommendation presented. The Committee discussed several other important considerations, including the establishment of a threshold percentage on what should be called a disparity and

# NATIONAL QUALITY FORUM

the concept of trends. In particular, the Committee raised a concern with the potential masking of disparity improvements—that is, if the gap continues to widen but the trend is down, then another level of analysis should be considered. More specifically, there may be improvements, but those specific improvements may not address the disparity. Committee members suggested calculating trends, in addition to relative/absolute and favorable/adverse. They also thought it was important to provide specific rates, remain consistent with the method utilized, and provide a narrative for explanation.

## ***Paired Comparisons versus Summary Statistics (Section 4d)***

The Committee discussed the concept of paired comparisons versus summary statistics. The paper noted that most summary measures of disparities lack directionality, and therefore great care must be taken before using them to track disparities. It recommended that pairwise comparisons using the historically advantaged group as the reference point should be checked to see if a positive finding from the summary statistic reflects superior care received by the disadvantaged group—and if so, then the context of that result and relevant policy goals must be explicitly considered.

The Committee expressed no preference for paired comparisons versus summary statistics and discussed the pros and cons associated with each. Several Committee members articulated the importance of simplicity, or ease at which the end-user is able to understand and use the information at hand. Paired comparisons were identified as generally more transparent and user-friendly than summary statistics. Committee members cautioned that, when using a summary statistic, the process of condensing measures may often include loss of important information or involve subjective decisions to achieve the summary. In contrast, a paired comparison/ratio can be more easily interpreted. It was noted that NQF does not specifically use summary statistics in measures, but rather endorses composite measures, which may raise similar issues.

## ***Normative Judgments About Disparity Measures (Section 4e)***

The paper discussed normative judgments in the selection of disparities-sensitive measures to report. The Committee discussed the issue very briefly, noting that further evaluation of the measure and reference point for which the normative judgment is based would need to be explained. The Committee agreed that normative judgments and inherent biases should be minimized, and, when used in reported measures, they should be mentioned and referenced appropriately.

## ***Risk Adjustment and Stratification (Section 4i)***

The paper also addressed a topic that is an increasingly important consideration in NQF's current measure evaluation process—that is, the pros and cons of risk adjustment versus stratification. The paper discussed exception reporting as well. The authors recommended that stratification by race/ethnicity and primary language should be performed when there are sufficient data to do so and that risk adjustment may be appropriate when performance is highly dependent on community factors beyond a provider's control. (Existing NQF policy is that risk models should not include race/ethnicity adjusters, which may mask disparities in quality of care.)

The Committee generally agreed with the recommendation of stratification by race/ethnicity and primary language when sufficient data exist; however, considerable discussion took place

# NATIONAL QUALITY FORUM

regarding the various implications of risk adjustment versus stratification. Several Committee members articulated that there may not be a “one-size-fits-all answer” and “context is important.” In addition, the implications of this decision on payment systems and policy changes were noted. The Committee explored risk adjustment on two levels: 1) regarding an overall measure of quality, should risk adjustment be applied for race/ethnicity and language, and 2) when identifying disparities, should risk adjustment be applied for socioeconomic status or other indicators? Regarding the first question, the Committee ultimately acknowledged that a one-size-fits-all approach might not be desirable and should be undertaken with extreme caution so as not to mask disparities. Regarding the second question, the Committee agreed with the report’s recommendation that performance reports stratified by race/ethnicity should not be risk adjusted for socioeconomic status or other contributory factors, and instead should be further stratified if the data permit.

## ***Interaction Effects (Section 4g)***

The paper discussed interaction effects, specifically noting that the most common disparity comparison is made within a single domain, such as differences among racial groups or ethnicities. It noted, however, that disparities may in some cases exist only for subsets of a particular racial/ethnic group; in statistical terms, this is referred to as an interaction effect—that is, the situation where the effect of one group differs depending on the characteristics (or level) of the other group. The paper thus recommended that, when clear differences in quality exist by racial/ethnic sub-strata, further stratification of results serves to highlight areas of the greatest potential for intervention.

The Committee generally agreed with the recommendation related to interaction effects and offered additional variables to consider for stratification, including income, age, highest level of education, acculturation, and urban/rural effects and language, to further elucidate areas for intervention. The rarity and “size” of an event that may contribute to an interaction effect is usually small; therefore caution should be exercised when using these small events to develop a conclusion with generalizing disparities. The Committee agreed that an interaction effect should be acknowledged, but reported only if it is large enough to make a difference on the disparity.

## ***Sample Size Consideration (Section 4h)***

The paper noted that identification of disparities is often hampered by sample sizes, because many racial/ethnic groups are in the minority. Disparities measurement and reporting programs (and incentive programs) face a major challenge when providers or institutions have small numbers of minority patients. The paper provided several options (e.g., rolling up data, summary statistics, use of composite measures, and combining data from two or more years) to consider when addressing the issue of small sample size, highlighting the pros and cons for each. The paper noted that the strategy deployed will be measure and context specific.

The Committee generally agreed with the options presented, except for the use of summary statistics, which generally consider all racial/ethnic groups simultaneously. As discussed in a previous section, summary statistics can be useful to reveal disparities but should not be used blindly. Therefore, the Committee did not recommend the wholesale use of summary statistics but suggested it remain as an option.

# NATIONAL QUALITY FORUM

The Committee discussed the pros and cons of the other options, including the feasibility of using composite measures with small sample sizes, the adequacy of sample size for measurement, the establishment of a sampling plan, and granularity when reporting small sample size. After considering the issues, the Committee recommended over sampling for race, ethnicity, and language as well as other sub-groups. The Committee also acknowledged the importance of collecting data regardless of the sample size: Data collected on a population or minority group may reveal something important or demonstrate a pattern for further research and may be effective for internal quality improvement, but may not be strong enough to publicly report or make generalizations. In this regard, it was noted that circumstantial issues may be relevant as it relates to sample size; statistical ability alone does not always tell the whole story.

## ***Consideration of Socioeconomic and Other Demographic Variables (Section 4i.ii)***

The paper discussed the use of socioeconomic and other demographic variables for stratification. The paper recommended that performance reports stratified by race/ethnicity should not be risk adjusted by socioeconomic status or other contributory factors, and instead should be further stratified if the data permit. The Committee generally agreed with this recommendation; however, a few concerns were noted and discussed. First, the Committee noted the sensitivity of over adjusting for disparities within a population and questioned whether it can be concluded that such disparities really don't exist if they go away once the measure is adjusted for variables such as socioeconomic status and income. The Committee also believed that it is important to differentiate this recommendation from other risk-adjustment activity—that is, risk adjusting an outcome or measure for race/ethnicity. Overall, the Committee concluded that stratification is a better option for pulling out differences in the underlying racial/ethnic population. The Committee also discussed the indicator of wealth versus income and noted it could be important to assess this difference.

## **DISCUSSION: PRIORITIES AND OPTIONS FOR PUBLIC REPORTING AND POLICY IMPLICATIONS**

The paper provided several recommendations for priorities and options for public reporting and policy implications; the Committee was asked to review these considerations and provide additional feedback.

Several suggestions were made including the concept of motivating providers to improve performance—the usage of a consumer liaison or “middle person” to serve as a mediator between the community and providers/organizations, the concept of incentivizing the patient toward better behavior modifications, and measurement of the effectiveness of services provided to the patient. Larger global issues were addressed such as the importance of making quality improvement more efficient, emphasizing that all providers should be addressing disparities to be considered a quality service provider, demonstrating the impact of reducing disparities and the tie to healthcare costs reduction, and improving the number of minorities utilized for clinical research.

In addition, the Committee discussed the concept of public reports as means to motivate providers. The Committee noted the difference between public reports for internal purposes and

# NATIONAL QUALITY FORUM

those for consumers, as well as the social benefits of reducing disparities, particularly for organizations providing care to large minority populations.

## **DISCUSSION: FRAMING THE CALL FOR MEASURES AND NQF APPROACH FOR DISPARITIES MEASUREMENT**

### **Framing the Call for Measures**

In order to insure that measures submitted to the Healthcare Disparities and Cultural Competency Consensus Standards project are appropriate and meaningful, Committee members were asked to provide input to the Call for Measures for this project's second phase. Ms. McElveen explained NQF's standard issuance of a Call for Measures, a step of the Consensus Development Process (CDP). The Committee identified the following issues and items as important for consideration:

- System/structural measures for capturing disparities
- Cross-cutting measures applicable for all populations
- Access
- Reducing ambulatory care disparities
- Health information technology
- Health-related quality of life
- Cultural competency
- Language services
- Integrating health communication (language, literacy, and culture)
- Resource allocation
- Patients' use of "navigation" devices
- Education for academic centers
- Adherence to core competencies
- Patient perceptions/bias
- Care coordination
- Diversity of leadership, staff, and community advisory groups
- Community engagement
- Highlighting the level of analysis
- Larger global issues for quality improvement (e.g., readmissions)
- Medical homes.

### **NQF's Approach for Measuring Disparities Prospectively**

The Committee provided several recommendations for NQF's approach for addressing disparities measurement both in the current form and prospectively. The Committee suggested that the entire NQF portfolio of measures be reviewed and a subset of disparities-sensitive measures identified. The selection criteria of prevalence and quality gap should serve as starting points for identifying those measures; following this initial review, the additional criteria would be applied for an additional filter. Once the starter set is identified, it will be important to convey



# NATIONAL QUALITY FORUM

or disseminate the criteria used to select it. This is of particular importance for organizations seeking to select measures for addressing disparities. The Committee also suggested that NQF consider process and outcome measures separately or look at system-based and provider-based measures first and then identify the cross-cutting measures. The Committee noted that first screening and then organizing the measures through a “protocol” may identify not only disparities gaps within the portfolio but also themes that are appropriate for new measure development. The Committee also agreed that it is necessary to stratify all measures by race/ethnicity and language and considered prioritizing measures for implementation and uptake by various institutions.

The Committee also briefly reviewed the NQF measure submission form to specifically identify possible improvements and changes around the current questions related to disparities. The Committee thought it was important for the disparities issue to be front and center on the submission form and offered several suggestions, including: advising measure developers more specifically about including disparities; aggregating the currently dispersed disparities sections within each evaluation criterion to a new, separate section toward the beginning of the form; and considering disparities as a threshold criterion for NQF endorsement. In addition, the Committee thought it was important to provide developers with a clear definition for disparities and disparities-sensitive measures and to reference the recommendations from the commissioned paper.

## **NEXT STEPS**

Ms. McElveen outlined the next steps that would occur after the meeting’s close. NQF staff will provide the Committee with several documents for review, including the meeting summary, the protocol to be used for review of NQF portfolio measures, the conclusions/recommendations document regarding the methodological issues that may serve as guidance to Steering Committees and/or measure developers, the draft Call for Measures, and a schedule for upcoming conference calls.

The next in-person Steering Committee meeting is currently scheduled for January 2012.