

Cancer, Fall 2018 Measure Review Cycle

Standing Committee Orientation Melissa Mariñelarena, Senior Director Katie Goodwin, Senior Project Manager Tara Murphy, Project Manager

December 12, 2018

Welcome

Project Team

- Melissa Mariñelarena, RN, MPA, CPHQ, Senior Director
- Katie Goodwin, MS, Senior Project Manager
- Tara Murphy, MPAP, Project Manager

Agenda for the Call

- Standing Committee introductions
- Overview of NQF, the Consensus Development Process, and Roles of the Standing Committee, co-chairs, NQF staff
- Overview of NQF's portfolio of cancer measures
- Review of project activities and timelines
- Overview of NQF's measure evaluation criteria
- Overview of Social Risk
- SharePoint Tutorial
- Measure worksheet example
- Next steps

Cancer Standing Committee

Karen Fields, MD, Co-Chair Shelley Fuld Nasso, MPP, Co-Chair Laurence Baker, DO Gregary Bocsi, DO, FCAP Brent Braveman, PhD, OTR/L, FAOTA Steven Chen, MD, MBA, FACS Matthew Facktor, MD, FACS Heidi Floyd Richard Gelb, MA Bradford Hirsch, MD Jette Hogenmiller, PhD, MN, APRN/ARNP, CDE, NTP, TNCC, CEE J. Leonard Lichtenfeld, MD, MACP

Robert Rosenberg, MD, FACR **Stephen Lovell** Jennifer Malin, MD, MACP Jodi Maranchie, MD, FACS Ali McBride, PharmD, MS, BCPS, BCOP Benjamin Movsas, MD Diane Otte, RN, MS, OCN Beverly Reigle, PhD, RN Robert Rosenberg, MD, FACR David J. Sher, MD, MPH Danielle Ziernicki, PharmD

Overview of NQF, the CDP, and Roles

The National Quality Forum: A Unique Role

Established in 1999, NQF is a nonprofit, nonpartisan, membership-based organization that brings together public and private sector stakeholders to reach consensus on healthcare performance measurement. The goal is to make healthcare in the U.S. better, safer, and more affordable.

Mission: To lead national collaboration to improve health and healthcare quality through measurement

- An Essential Forum
- Gold Standard for Quality Measurement
- Leadership in Quality

2 3 4

NQF Activities in Multiple Measurement Areas

Performance Measure Endorsement

- 600+ NQF-endorsed measures across multiple clinical areas
- **1**5 empaneled standing expert committees

Measure Applications Partnership (MAP)

Advises HHS on selecting measures for 20+ federal programs

National Quality Partners

- Convenes stakeholders around critical health and healthcare topics
- Spurs action: recent examples include antibiotic stewardship, advanced illness care, shared decision making, and opioid stewardship

Measurement Science

- Convenes private and public sector leaders to reach consensus on complex issues in healthcare performance measurement
 - » Examples include HCBS, rural issues, telehealth, interoperability, attribution, riskadjustment for social risk factors, diagnostic accuracy, disparities

Measure Incubator

 Facilitates efficient measure development and testing through collaboration and partnership

NQF Consensus Development Process (CDP) 6 Steps for Measure Endorsement

- Intent to Submit
- Call for Nominations
- Measure Evaluation
- Public Commenting Period with Member Support
- Measure Endorsement
- Measure Appeals

Measure Review: Two Cycles Per Year

Consensus Development Process:

Two Cycles Every Contract Year



15 New Measure Review Topical Areas



Denotes expanded topic area

^A Cost & Efficiency will include efficiency-focused measures from other domains

^B Geriatric & Palliative Care includes pain-focused measures from other domains

 $^{\rm C}$ Patient Safety will include acute infectious disease and critical measures

^D Prevention and Population Health is formerly Health and Well Being

Role of the Standing Committee General Duties

- Act as a proxy for the NQF multistakeholder membership
- Serve 2-year or 3-year terms
- 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

Role of the Standing Committee Measure Evaluation Duties

- All members evaluate ALL measures being considered for endorsement
- 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
- Oversee cancer portfolio of measures
 - Promote alignment and harmonization
 - Identify gaps

Role of the Standing Committee Co-Chairs

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

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 SC through the CDP and advise on NQF policy and procedures
 - Review measure submissions and prepare materials for Committee review
 - Draft and edit reports for SC review
 - Ensure and facilitate communication among all project participants (including SC and measure developers)
 - Facilitate collaboration between different NQF projects

Role of NQF Staff *Communication*

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

Role of Methods Panel

- Scientific Methods Panel created to ensure higher-level and more consistent reviews of the scientific acceptability of measures
- The Methods Panel is charged with:
 - Conducting evaluation of complex measures for the Scientific Acceptability criterion, with a focus on reliability and validity analyses and results
 - Serve in advisory capacity to NQF on methodologic issues, including those related to measure testing, risk adjustment, and measurement approaches.
- The Methods Panel review will help inform the Standing Committee's endorsement decision. The Panel will not render endorsement recommendations.

NQF Consensus Development Process (CDP) Measure Evaluation

Complex Measures	 Outcome measures, including intermediate clinical outcomes Instrument-based measures (e.g., PRO-PMs) Cost/resource use measures Efficiency measures (those combining concepts of resource use and quality) Composite measures
Noncomplex Measures	 Process measures Structural measures Previously endorsed complex measures with no changes/updates to the specifications or testing

Questions?

Overview of NQF's Cancer Portfolio

NATIONAL QUALITY FORUM

Cancer Portfolio of Measures

- This project will evaluate measures related to cancer that can be used for accountability and public reporting for all populations and in all settings of care. This project will address topic areas including:
 - Pain intensity for patients receiving chemotherapy or radiation therapy
 - Treatment of osteopenia or osteoporosis in men receiving androgen deprivation therapy for non-metastatic prostate cancer
 - Admissions and emergency visits for patients receiving outpatient chemotherapy
- NQF solicits new measures for possible endorsement
- NQF currently has 26 endorsed measures within this topic area. Endorsed measures undergo periodic evaluation to maintain endorsement – "maintenance".

Cancer Portfolio of NQF-endorsed measures

*Measures for maintenance evaluation

Breast

- 0508 Diagnostic Imaging: Inappropriate Use of "Probably Benign" Assessment Category in Screening Mammograms
- 0509 Diagnostic Imaging: Reminder System for Screening Mammograms
- **1878** Human epidermal growth factor receptor 2 (HER2) testing in breast cancer
- **0219** Post breast conservation surgery irradiation
- 1857 Patients with breast cancer and negative or undocumented human epidermal growth factor receptor 2 (HER2) status who are spared treatment with trastuzumab
- 1858 Trastuzumab administered to patients with AJCC stage I (T1c) – III and human epidermal growth factor receptor 2 (HER2) positive breast cancer who receive adjuvant chemotherapy
- 0387 Oncology: Hormonal therapy for stage IC through IIIC, ER/PR positive breast cancer
- 0559 Combination chemotherapy is considered or administered within 4 months (120 days) of diagnosis for women under 70 with AJCC T1cN0M0, or Stage IB
 III hormone receptor negative breast cancer
- **0220** Adjuvant hormonal therapy

Colon

- 0225 At least 12 regional lymph nodes are removed and pathologically examined for resected colon cancer
- 0385 Oncology: Chemotherapy for AJCC Stage III Colon Cancer Patients
- 0223 Adjuvant chemotherapy is considered or administered within 4 months (120 days) of diagnosis to patients under the age of 80 with AJCC III (lymph node positive) colon cancer
- 1859 KRAS gene mutation testing performed for patients with metastatic colorectal cancer who receive anti-epidermal growth factor receptor monoclonal antibody therapy
- 1860 Patients with metastatic colorectal cancer and KRAS gene mutation spared treatment with anti-epidermal growth factor receptor monoclonal antibodies

Cancer Portfolio of NQF-endorsed measures

*Measures for maintenance evaluation

Hematology

- 0377 Hematology: Myelodysplastic Syndrome (MDS) and Acute Leukemias: Baseline Cytogenetic Testing Performed on Bone Marrow
- 0379 Hematology: Chronic Lymphocytic Leukemia (CLL): Baseline Flow Cytometry
- 0378 Hematology: Myelodysplastic Syndrome (MDS): Documentation of Iron Stores in Patients Receiving Erythropoietin Therapy
- 0380 Hematology: Multiple Myeloma: Treatment with Bisphosphonates

Lung/Thoracic

1854 Barrett's Esophagus

Prostate

- 1853 Radical Prostatectomy Pathology Reporting
- 0389 Prostate Cancer: Avoidance of Overuse of Bone Scan for Staging Low Risk Prostate Cancer Patients
- 0390 Prostate Cancer: Adjuvant Hormonal Therapy for High Risk Prostate Cancer Patients

Other

- 0383 Oncology: Plan of Care for Pain Medical Oncology and Radiation Oncology (paired with 0384)
- *0384 Oncology: Pain Intensity Quantified Medical Oncology and Radiation Oncology (paired with 0383)
- **0386** Oncology: Cancer Stage Documented
- 2930 Febrile Neutropenia Risk Assessment Prior to Chemotherapy

Activities and Timeline *All times ET

Meeting	Date/Time
Committee Orientation Meeting	Wednesday, December 12, 2018, 11am-1pm ET
Committee Measure Evaluation Web Meeting	Friday, February 8, 2019, 2-4pm ET
Committee Measure Evaluation Web	Tuesday, February 12, 2019, 1-3pm ET
Meeting	
Committee Post-Measure Evaluation	Friday, February 15, 2019, 2-4pm ET
Web Meeting	
Committee Post-Comment Web	Wednesday, May 7, 2019, 2-4pm ET
Meeting	

Questions?

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 28-29 in the SC Guidebook)

- 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 (must-pass for maintenance measures): 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 31-39)*

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 32-38)

- 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 patientreported structure/process measures.

Rating Evidence: Algorithm #1 (page 35)

[Screen share Evidence algorithm]

Criteria emphasis is different for <u>new</u> vs. <u>maintenance</u> 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 evidence from last evaluation; Standing Committee to affirm no change in evidence IF changes in evidence, 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 (pages 40 – 50)

Extent to which the measure, <u>as specified</u>, 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 41)

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 42)*

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 43)

- 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 44)

[Screen share Reliability algorithm]

Validity testing (pages 45-49)

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 49)

[Screen share Validity algorithm]

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 (pages 50-51)

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 process3b: Electronic sources3c: Data collection strategy can be implemented

Criterion #4: Usability and Use (pages 51-52)

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) 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

Feasibility

New measures	Maintenance measures		
 Measure feasible, including eCQM feasibility assessment 	NO DIFFERENCE: Implementation issues may be more prominent		

Usability and Use

New measures	Maintenance measures	
• Use: used in accountability applications and public reporting	INCREASED EMPHASIS : Much greater focus on measure use and	
 Usability: impact and unintended consequences 	usefulness, including both impact and unintended consequences	

Criterion #5: Related or Competing Measures (pages 52-53)

If a measure meets the four criteria <u>and</u> there are endorsed/new **related** measures (same measure focus or same target population) or **competing** measures (both the same measure focus <u>and</u> 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

- For CY2019 and beyond, reliability testing should be based on ICD-10 coded data.
- Validity testing should be based on ICD-10 coded data
- If providing face validity (FV), both FV of the ICD-10 coding scheme and FV of the measure score as an indicator of quality is required update

eCQMs

"Legacy" eCQMs

 Beginning September 30, 2017 all respecified measure submissions for use in federal programs will be required to the same evaluation criteria as respecified measures – the "BONNIE testing only" option will no longer meet endorsement criteria

For all eCQMs: Reliance on data from structured data fields is expected; otherwise, unstructured data must be shown to be both reliable and valid

Evaluation Process

- Preliminary analysis (PA): To assist the Committee evaluation of each measure against the criteria, NQF staff and 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 will conduct an in-depth evaluation on all measures under review
 - 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 inperson/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 comments submitted
- Final endorsement decision by the CSAC
- Appeals (if any)

Questions?

Social Risk Overview

NATIONAL QUALITY FORUM

Background

- NQF conducted a two-year trial period from 2015-2017. During this time, adjustment of measures for social risk factors was no longer prohibited
- The NQF Board of Directors reviewed the results of the trial period and determined there was a need to launch a new social risk initiative
- As part of the Equity Program, NQF will continue to explore the need to adjust for social risk
- Each measure must be assessed individually to determine if SDS adjustment is appropriate (included as part of validity subcriterion)
- The Standing Committee will continue to evaluate the measure as a whole, including the appropriateness of the risk adjustment approach used by the measure developer
- Efforts to implement SDS adjustment may be constrained by data limitations and data collection burden

Standing Committee Evaluation

- The Standing Committee will be asked to consider the following questions:
 - Is there a conceptual relationship between the SDS factor and the measure focus?
 - What are the patient-level sociodemographic variables that were available and analyzed during measure development?
 - Does empirical analysis (as provided by the measure developer) show that the SDS factor has a significant and unique effect on the outcome in question?
 - Does the reliability and validity testing match the final measure specifications?

Questions?

http://share.qualityforum.org/Projects/cancer/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

Screen shot of homepage:



Please keep in mind:

+ and – signs :

	Measure Documents		Measure Documents		
	Measure Number	Name	Measure Numbe	r Name	Description
~	Measure Sub-Topic: (1)		easure Sub-Topic: (1)		
	Add document		0521	Heart Failure Symptoms Assessed and Addressed	Percentage of home health episodes heart failure were assessed for sym appropriate actions were taken whe heart failure.
			🕈 Add document		
	Meeting and Call Documents	Meeting and Call Documents			
	Type Nar	ne	Type Na	ame	
₽	Meeting Title : 1/30/2014 Orientation Call (1) Add document		Meeting Title : 1/30/2014 Orientation Call (1) NQF Cardiovascular Project Orientation Agenda		
-			Add document		

Measure Worksheet and Measure Information

Measure Worksheet

- 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

Measure Evaluation Web Meetings

- Friday, February 8, 2019, 2-4pm ET
- Tuesday, February 12, 2019, 2-4pm ET
- Friday, February 15, 2019, 2-4pm ET

Project Contact Info

- Email: <u>cancerm@qualityforum.org</u>
- NQF phone: 202-783-1300
- Project page: <u>http://www.qualityforum.org/Project_Pages/cancer.aspx</u>
- SharePoint site: <u>http://share.qualityforum.org/Projects/cancer/SitePages</u> /<u>Home.aspx</u>

Questions?

