NATIONAL QUALITY FORUM

Measure Sets and Measurement Systems

Measure Systems Subgroup Meeting #1
July 17, 2019

Agenda

- Welcome
- Roll Call
- Objectives of TEP and Today's Meeting
- Orientation Meeting Recap
- Draft Measure Systems Key Elements Discussion
  - *Intent*
    - Purpose
    - Quality construct
  - *Measure Set*
    - Specification alignment
    - Domains/ grouping
- Next Steps

NQF Project Staff

- Measure Systems Subgroup
  - *Nicolette Mehas, PharmD, Director*
  - *Taroon Amin, PhD, Consultant*
  - *Madison Jung, Project Manager*
  - *Asaba Mbenwoh Nguafor, RN, MSN/MPH, Project Analyst*

Measure Sets and Measurement Systems Technical Expert Panel

Measurement Systems Subgroup

- **Co-chair: Michael Chernew, PhD**
- **Co-chair: Sam Simon, PhD**
- Philip Alberti, PhD
- Matt Austin, PhD
- Kari Baldonado
- Julie Bershadsky, PhD
- Amy Chin, MS
- Melissa Danforth
- Marybeth Farquhar, PhD, MSN, RN
- Frank A. Ghinassi, PhD, ABPP
- Danielle Lloyd, MPH
- Jeffrey Sussman, MPH, PhD

Project Motivation

- NQF convenes healthcare stakeholders to evaluate the scientific merits of a performance measure and to provide guidance on its use, but there is **no current process for evaluating how measures work together**.

- There is a **lack of consistency, transparency, and stakeholder input** in the way that individual performance measures are used together, either as part of a measure set or a larger measurement system, such as an APM.

Project Goal

- Objectives
  - *Build upon and refine the key components of a measurement system.*
  - *Discuss and come to consensus on best practices, principles, and criteria (if possible) that can be used for the development and evaluation of a measurement system.*

- Results
  - *Draft and final report on best practices and principles for measure sets and measurement systems*
    - » This report is the first step in determining how to systematically and transparently evaluate measure sets and measurement systems.

Measurement Systems Recap from Orientation Meeting

Draft Definition

Measurement System – a group of individual measures that, based on a predefined methodology, work together to assess quality or cost in relationship to a goal

Components to consider:

- Intent (i.e., purpose and quality construct)
- Measure Set (e.g., specification alignment, domains)
- Aggregation
- Incentive mechanism
- Attribution methodology
- Risk adjustment

← *Focus of today's meeting*

Differences between Sets and Systems

- Measure sets are NOT aggregated to create a single composite score.
- A measure set PLUS other programmatic elements: aggregation, incentive mechanism, risk adjustment, and attribution model forms a measurement system.

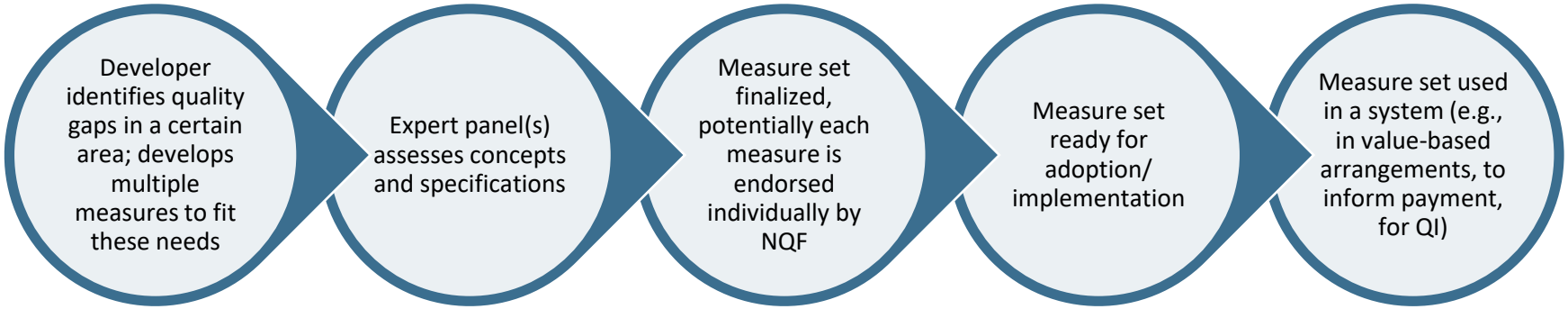
Steps in Creating a Measure Set (from White Paper)

1. Identifying a purpose
2. Defining a quality construct
3. Selecting measures to assess that quality construct

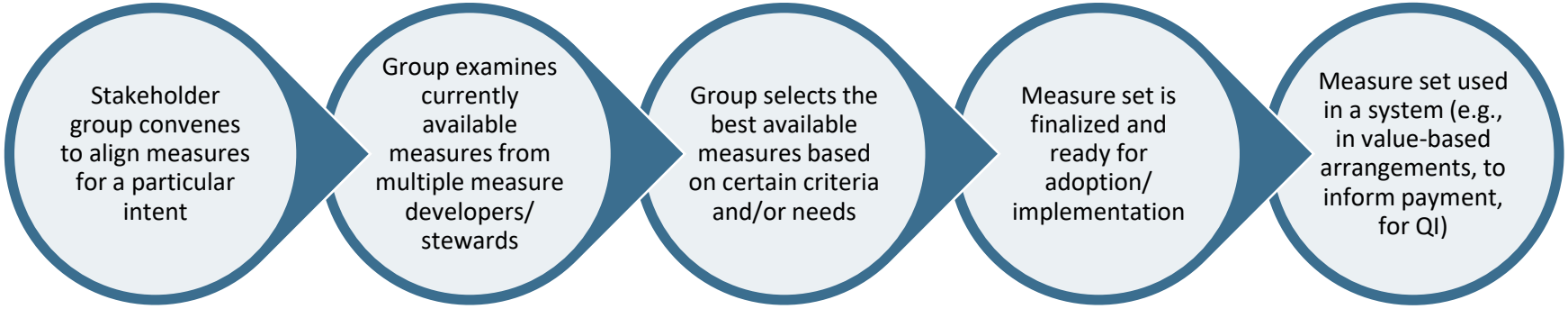
These steps are also the beginning of measure system development.

Measure Set Development Pathways

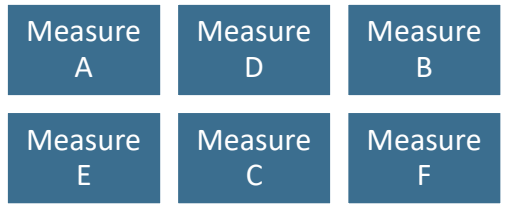
Internal development of measures to be used together to best capture quality for a certain purpose



External group creates a measure set with the goal of aligning measure use

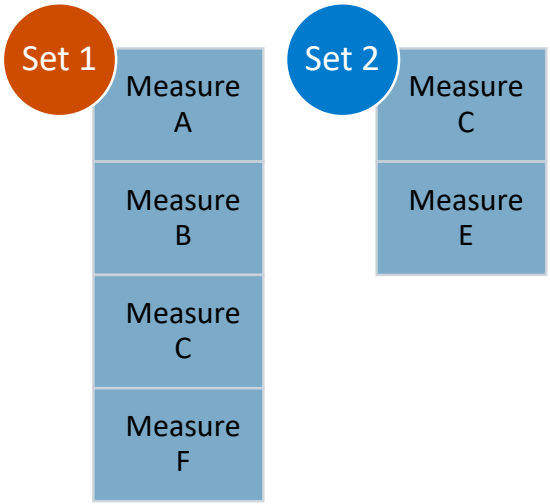


Conceptual Framework: Quality Measurement Infrastructure



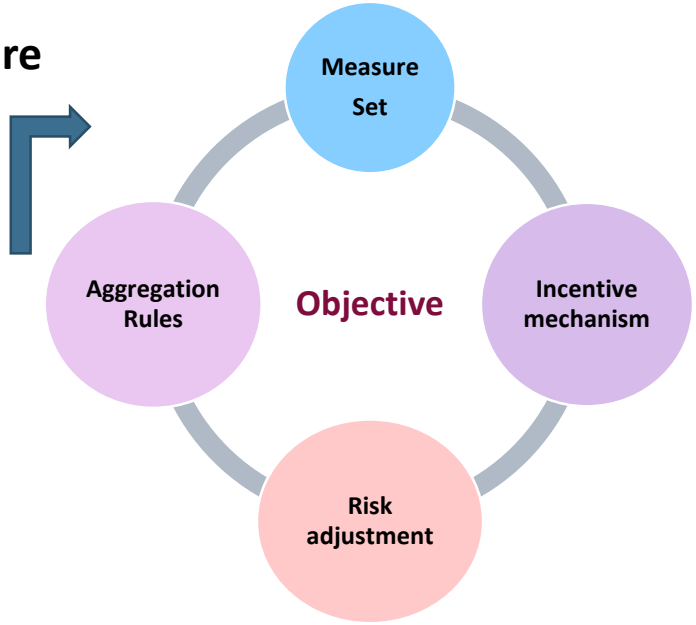
Level 1: Individual Measures

Example: NQF Measure 2158 Medicare Spending Per Beneficiary



Level 2: Measure Sets

Example: CMS Hospital Value-Based Purchasing Program (HVBP) uses a set of measures assessing four domains: safety, clinical care, efficiency and cost reduction, and patient experience.



Level 3: Measurement Systems

Example: Measures are aggregated as part of the HVBP program, and the score is used to determine hospital penalties or rewards.

Feedback from TEP Orientation

- Consistent definitions are important…what other definitions currently exist?
- “Intended use” must be specific and thoughtful
- Consider level of attribution when constructing a measure set
- Proactive approach to systems (users should be aware of measures beforehand)
- Adaptive approach to systems and “pick-list” topics need further discussion
- Consider timing and cycle frequency as additional components
- Consider how key components interact with one another and the burden versus benefit of approaches

Key Elements Discussion: *Intent & Measure Set*

Intent - Purpose

- Elements:
  - *Focus (e.g., cardiology, gastroenterology, safety, primary care)*
  - *Intended Use and User (e.g., for use by plans for VBP, by states for benchmarking, to support patient choice of providers, for QI)*
  - *Target population*
  - *Accountable entity (e.g., clinicians, hospitals, ACOs)*
  - *Policy or programmatic goals*
- Example draft recommendations:
  - *A measure set should clearly specify the 1) focus, 2) target population, 3) accountable entity, 4) intended use and user.*
  - *The intended use of a measure set should be explainable to those being measured and using performance results.*
- Questions:
  - *Are all of these sub-components important to specify? Any missing?*
    - » Is setting important to specify? Is focus important to specify?
  - *Key considerations when evaluating this element?*
    - » Should we be general or specific with guidance related to "purpose"?

Background: Approaches for Constructing Composites (related to quality construct)

1. The quality construct is seen as causing or reflected in the component measure scores.
2. The quality construct is seen as being caused or defined by the component measure scores.
3. The quality construct is viewed or defined as receiving all necessary care represented by the component measures (all or none or partial credit)
4. The quality construct is viewed as individual patients not experiencing any healthcare-acquired adverse event/complication or not receiving unnecessary or inappropriate care (any or none)

From NQF Composite Performance Measure Evaluation Guidance, Appendix B

Background: NQF Composite Measure Requirements (related to quality construct)

Importance to Measure and Report

1c. For composite performance measures, the following must be explicitly articulated and logical

- **1c1.** The quality construct, including the overall area of quality; included component measures; and the relationship of the component measures to the overall composite and to each other; and
- **1c2.** The rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually; and
- **1c3.** How the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.

From NQF Measure Evaluation Criteria

Background: NQF Composite Measure Requirements (related to quality construct)

Scientific Acceptability of Measure Properties

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate the following:

- **2c1.** the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and
- **2c2.** the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

If not conducted or results not adequate, justification must be submitted and accepted

From NQF Measure Evaluation Criteria

Intent - Quality Construct

- Elements:
  - *Overall area of quality; Included "component" measures; Conceptual relationships between each component and the overall "composite"; Relationships among the "component" measures*
- Considerations:
  - *Conceptual model*
  - *Evidence base (impact/prevalence/national priorities/guidelines?)*
  - *Parsimony*
  - *Construct for cost/efficiency*
- Example draft recommendation (needs expansion):
  - *A measure set should be based on an evidence-based quality construct that relates to its predefined purpose.*
- Questions:
  - *Are these subcomponents important? Any missing?*
  - *Key considerations when evaluating this element? What testing should be used for "quality construct"?*
  - *How specific can we be with guidance? How do NQF's criteria for composite measures (slide 16 and 17) relate here?*

Background: MAP

- The Measure Applications Partnership (MAP) is a public-private partnership convened by NQF to inform the selection of performance measures in federal programs.

- MAP brings together stakeholders in a unique collaboration that balances the interests of consumers, purchasers, labor, health plans, clinicians and providers, communities and states, and suppliers.

- MAP uses its Measure Selection Criteria to guide its review of measures under consideration.
  - *The Measure Selection Criteria are intended to assist MAP with identifying characteristics that are associated with ideal measure sets used for public reporting and payment programs.*

Background: MAP Measure Selection Criteria

1. NQF-endorsed measures are required for program measure sets, unless no relevant endorsed measures are available to achieve a critical program objective

Demonstrated by a program measure set that contains measures that meet the NQF endorsement criteria, including importance to measure and report, scientific acceptability of measure properties, feasibility, usability and use, and harmonization of competing and related measures

- **Subcriterion 1.1** Measures that are not NQF-endorsed should be submitted for endorsement if selected to meet a specific program need
- **Subcriterion 1.2** Measures that have had endorsement removed or have been submitted for endorsement and were not endorsed should be removed from programs
- **Subcriterion 1.3** Measures that are in reserve status (i.e., topped out) should be considered for removal from programs

2. Program measure set actively promotes key healthcare improvement priorities, such as those highlighted in CMS' "Meaningful Measures" Framework

Demonstrated by a program measure set that promotes improvement in key national healthcare priorities such as CMS's Meaningful Measures Framework.

- Other potential considerations include addressing emerging public health concerns and ensuring the set addresses key improvement priorities for all providers.

Background: MAP Measure Selection Criteria

3. Program measure set is responsive to specific program goals and requirements

Demonstrated by a program measure set that is "fit for purpose" for the particular program

- **Subcriterion 3.1** Program measure set includes measures that are applicable to and appropriately tested for the program's intended care setting(s), level(s) of analysis, and population(s)
- **Subcriterion 3.2** Measure sets for public reporting programs should be meaningful for consumers and purchasers
- **Subcriterion 3.3** Measure sets for payment incentive programs should contain measures for which there is broad experience demonstrating usability and usefulness (Note: For some Medicare payment programs, statute requires that measures must first be implemented in a public reporting program for a designated period)
- **Subcriterion 3.4** Avoid selection of measures that are likely to create significant adverse consequences when used in a specific program
- **Subcriterion 3.5** Emphasize inclusion of endorsed measures that have eMeasure specifications available

Background: MAP Measure Selection Criteria

4. Program measure set includes an appropriate mix of measure types

Demonstrated by a program measure set that includes an appropriate mix of process, outcome, experience of care, cost/resource use/appropriateness, composite, and structural measures necessary for the specific program

- **Subcriterion 4.1** In general, preference should be given to measure types that address specific program needs
- **Subcriterion 4.2** Public reporting program measure sets should emphasize outcomes that matter to patients, including patient- and caregiver-reported outcomes
- **Subcriterion 4.3** Payment program measure sets should include outcome measures and cost measures to capture value

5. Program measure set enables measurement of person- and family-centered care and services

Demonstrated by a program measure set that addresses access, choice, self-determination, and community integration

- **Subcriterion 5.1** Measure set addresses patient/family/caregiver experience, including aspects of communication and care coordination
- **Subcriterion 5.2** Measure set addresses shared decision-making, such as for care and service planning and establishing advance directives
- **Subcriterion 5.3** Measure set enables assessment of the person's care and services across providers, settings, and time

Background: MAP Measure Selection Criteria

6. Program measure set includes considerations for healthcare disparities and cultural competency

Demonstrated by a program measure set that promotes equitable access and treatment by considering healthcare disparities. Factors include addressing race, ethnicity, socioeconomic status, language, gender, sexual orientation, age, or geographical considerations (e.g., urban vs. rural). Program measure set also can address populations at risk for healthcare disparities (e.g., people with behavioral/mental illness).

- **Subcriterion 6.1** Program measure set includes measures that directly assess healthcare disparities (e.g., interpreter services)
- **Subcriterion 6.2** Program measure set includes measures that are sensitive to disparities measurement (e.g., beta blocker treatment after a heart attack), and that facilitate stratification of results to better understand differences among vulnerable populations

7. Program measure set promotes parsimony and alignment

Demonstrated by a program measure set that supports efficient use of resources for data collection and reporting, and supports alignment across programs. The program measure set should balance the degree of effort associated with measurement and its opportunity to improve quality.

- **Subcriterion 7.1** Program measure set demonstrates efficiency (i.e., minimum number of measures and the least burdensome measures that achieve program goals)
- **Subcriterion 7.2** Program measure set places strong emphasis on measures that can be used across multiple programs or applications

Background: CQMC

- Partnership between AHIP, CMS, NQF and approximately 60 member organizations including medical associations, health insurance providers, consumers, purchasers, etc.

- Goal: To align measures used by all payers, both public and private, in various specialty and other topics areas (e.g., primary care, pediatrics)

- The CQMC's Selection Principles allow members to weigh the merits of an individual measure and determine if a core set is comprehensive and aligned with the CQMC's vision.
  - *The selection principles consider various stakeholder priorities and aim to balance valued concepts.*

Background: CQMC Selection Principles

Principles for measures included in the CQMC core measure sets

- Advance health and healthcare improvement goals and align with stakeholder priorities.
  - *Address a high-impact aspect of healthcare where a variation in clinical care and opportunity for improvement exist.*
- Are unlikely to promote unintended adverse consequences.
- Are scientifically sound (e.g., NQF-endorsed or otherwise proven to be evidence-based, reliable, and valid in diverse populations).
  - *The source of the evidence used to form the basis of the measure is clearly defined.*
  - *There is high quality, quantity, and consistency of evidence.*
  - *Measure specifications are clearly defined.*
- Represent a meaningful balance between measurement burden and innovation.
  - *Minimize data collection and reporting burden, while maintaining clinical credibility (i.e., measures that fit into existing workflows, are feasible, and do not duplicate efforts).*
  - *Are ambitious, yet providers being measured can meaningfully influence the outcome and are implemented at the intended level of attribution.*
  - *Are appropriately risk adjusted and account for factors beyond control of providers, as necessary.*

Background: CQMC Selection Principles

Principles for the CQMC core measure sets

- Provide a person-centered and holistic view of quality, including consideration of Social Determinants of Health (SDOH) and experience of care.
- Provide meaningful and usable information to all stakeholders.
- Promote parsimony, alignment, and efficiency of measurement (i.e., minimum number of measures and the least burdensome measures).
- Include an appropriate mix of measure types while emphasizing outcome measures and measures that address cross-cutting domains of quality.
- Promote the use of innovative measures (e.g., eMeasures, measures intended to address disparities in care, or patient-reported outcome measures).
- Include measures relevant to the medical condition of focus (i.e., “specialty-specific measures”).

Measure Selection – Specification Alignment

- Questions
  - *What are the necessary consistencies across a measure set?*
    - » Level of analysis (accountable individual or entity), setting, data source, denominator population
  - *Is this based on specific intent or are there "rules" for specification consistencies that can be generalized for all sets?*
  - *Do risk adjustment considerations apply at the time of measure selection?*
  - *How would you edit the below example statements?*

- Example draft recommendations:
  - *All measures in a set should be specified and tested at the appropriate level of analysis and in a population that represents the intended use.*
  - *All measures in a set should be linked by consistencies in accountable entity, setting, and/or population.*

Measure Selection – Domains (Groups)

- Definition/Examples
  - How measures within a set are arranged in subgroups; domains in a sets may be weighted differently.
  - MAP recommends measures for the HVBPP, but doesn't comment on what domain a measure should be in, what other measure should be in a domain, or how domains are weighted in the final scoring algorithm.
  - In the Overall Hospital Quality Star Rating program, measures within a group are weighted differently and group scores contribute differently to the overall score.

- Example draft recommendations:
  - *Special attention should be paid to how individual measures are assigned to domains and if a domain is both parsimonious and comprehensive.*
  - *Multiple stakeholders should provide guidance on how measures are assigned to domains, especially if domains are weighted differently.*

- Questions:
  - *How should grouping be evaluated? How can the above statements be more specific? How does one determine is grouping is appropriate? What are the methodological options for testing domains?*

Next Steps

Objectives for Future Meetings

- Subgroup Meetings #2 and 3 will focus on other key components: aggregation, incentive mechanism attribution methodology, and risk adjustment

- Subgroup Meeting #4 will focus on discussing criteria that emerged throughout the discussions, additional components and considerations identified by the group, and potential unintended consequences.

Upcoming Meeting Schedule

Meeting	Date/Time
Measurement Systems Subgroup Meeting #1	July 17, 2019 at 12:00PM-2:00 PM ET
Measure Sets Subgroup Meeting #2	August 14, 2019 at 12:00PM-2:00 PM ET
Measurement Systems Subgroup Meeting #2	August 20, 2019 at 12:00PM-2:00 PM ET
Full TEP Meeting	September 18, 2019 at 3:00-5:00 PM ET
Measurement Systems Subgroup Meeting #3	September 26, 2019 at 12:00PM-2:00 PM ET
Full TEP Meeting	October 30, 2019 at 12:00PM-2:00 PM ET
Measure Sets Subgroup Meeting #3	November 21, 2019 at 12:00PM-2:00 PM ET
Measurement Systems Subgroup Meeting #4	March 17, 2020 at 12:00PM-2:00 PM ET

Project Contact Info

- Email: measuresets@qualityforum.org
- NQF phone: 202-783-1300
- Project page: http://www.qualityforum.org/Measure_Sets_and_Measurement_Systems.aspx
- SharePoint site: http://share.qualityforum.org/Projects/MeasureSets/SitePages/Home.aspx