

NATIONAL QUALITY FORUM

IN-PERSON MEETING FOR RESOURCE USE STEERING COMMITTEE

July 12-13, 2010

Committee Members Present: Doris Lotz, MD, MPH (Co-Chair); Bruce Steinwald, MBA (Co-Chair); Paul Barnett, PhD; Jack Bowhan; Jephtha Curtis, MD; Kurtis Elward, MD, MPH; William Golden, MD; Lisa Grabert, MPH; Ethan Halm, MD, MPH; Ann Hendrich, RN, PhD(c); Thomas Lee, MD; Renee Markus-Hodin, JD; Jack Needleman, PhD; Mary-Kay O’Neill, MD, MBA; David Penson, MD, MPH; Steve Phillips, MPA; David Redfearn, PhD; Jeffrey Rich, MD; William Rich, MD; Barbara Rudolph, PhD, MSSW; Joseph Stephansky, PhD; Dolores Yanagihara, MPH; Tom Rosenthal, MD; James Weinstein, DO, MS

NQF Staff Present: Janet Corrigan, PhD, MBA; Helen Burstin, MD, MPH; Marybeth Farquhar, PhD; Heidi Bossley, MSN, MBA; Sally Turbyville, MA, MS; Jennifer Podulka, MPAff; Ashlie Wilbon, RN, MPH; Maisha Mims, MPH

Other Attendees: Niall Brennan (CMS), Rita Munley Gallagher, Kate Goodrich

WELCOME AND INTRODUCTIONS

Ms. Turbyville welcomed the Resource Use Steering Committee members and reviewed the agenda for the in-person two-day meetings. Ms Wilbon gave an overview of the Resource Use project activities to date as well as future project activities to come.

DAY 1: MONDAY, JULY 12, 2010

Meeting Objectives

- Resource use measure development and the environment
- Resource use white paper discussion
 - Section 1. Introduction and Reason for white paper
 - Defining resource use and efficiency
 - Section 2. Designing Measures That Acknowledge the Real World
 - Resource use measure continuum mode
 - Section 3. Types of Resource Use Measures
 - Data preparation
 - Creation of units for measurement
 - Applying units in measurement
 - Section 4. Limitations, Implications, and Open Questions

Resource Use Measure Development and the Environment

The Co-Chairs invited a representative of the project funder (CMS), Niall Brennan, to provide a brief summary of CMS’s goals related to this project. The Committee was given the opportunity to ask questions to better understand the implications of this project on CMS’s goals and the impact of resource use measurement in the healthcare environment, particularly for providers.

- Mr. Brennan stated that this project was conceived and began prior to the requirement for a public domain grouper came into law; however, this project and the Request for Proposals are complementary.

NATIONAL QUALITY FORUM

- This project's primary goal is to establish consensus and standards around evaluating resource use measures. The findings from this project can feed the public domain grouper process.
- The award for the public grouper is expected to be made in the early fall of this year. The grouper is expected to be complete by January 2012. CMS has approached this effort with a short- and long-term goal that is, adapting existing methodologies for the short-term and potentially creating a new methodology for the long-term with more time and resources.
- A Steering Committee member asked if this project also complements the efforts of the Affordable Care Act (ACA) as it relates to transitioning the healthcare system from fee-for-service to an integrated delivery system.
 - Mr. Brennan agreed and acknowledged that the move to value-based purchasing is a key driver of health reform. Within this structure, accountable care organizations (ACOs) will be evaluated using quality and cost-of-care metrics, and physicians will be evaluated on the basis of cost or quality under the value modifier provision. He went on to explain that the process for defining the physician value modifier is still some time away. The measures evaluated in this project are viewed as "first-generation" resource use measures.
- The Committee sought to clarify how the short-term goal of identifying individual measures for endorsement through this project might impact the long-term goals of identifying value for individual physicians and moving away from fee-for-service to ACOs. Would the decisions and efforts of this Committee potentially limit future decisions on measurement options (i.e., data organization, collection, reporting, etc.) and how people use measures?
- NQF endorses measures for three years, recognizing that time and experience changes what we know about measurement. As a voluntary consensus-based organization, CMS should look first to NQF to determine which consensus standards should be used in a program. CMS, however, is not the only user of NQF-endorsed® measures; purchasers, communities, and others also look to NQF for endorsed measures.

Resource Use Measurement White Paper Discussion

The Co-Chairs led the Committee members through the various sections of the white paper.

Section 1: Defining resource use and efficiency

Resource use. Tasked with defining resource use, the Committee developed the following definition for resource use:

- Resource use measures should accurately count what resources have been used by a delivery system (e.g., MD, hospital, etc.) for a specific population. They should count first, and monetize second (standard vs. actual).
 1. Counts of resources
 2. Monetize resources
 - Can be a standard cost approach or allowable charges, etc.
 - Allows for aggregation across resources
 3. Counting and monetizing is done and applied to a population or a sample of a population (denominator)
- Resource use measures are a building block to the value equation.
- Resource use measures are broadly applicable, and comparable measures of input counts (in terms of units or dollars) can be in terms of relative value units or specific services.
- The Committee suggested that the first thing to count should be inputs on a given population or sample of population.

NATIONAL QUALITY FORUM

- The white paper should explicitly state whether counting resources should be done in relation to improving health outcomes. The purpose of counting resources is to figure out how to count inputs to relatively rank as a return on investment for quality outcomes.

Efficiency. The committee agreed that explaining the difference between efficiency and resource use would be best illustrated through an equation between efficiency, quality, and resource use:

Efficiency = (fx)(resource use & quality)

Section 2: Designing measures that acknowledge the real world

Resource use measure continuum model



- The Committee agreed that additional information may be needed to clarify the differences between each resource use measurement approach on the continuum model.
- Some Committee members questioned the difference between per capita and per patient episodes. Committee members stated that there may be more value in focusing on measures on the per capita end of the spectrum. However, risk-adjustment is often very complex in per capita measures and could be more divisive.
- A Committee member suggested that per condition, per hospitalization, and per event be added to the continuum model.
- A Committee member noted that this continuum intersects with attribution. Attribution becomes somewhat simpler on the right end of the spectrum as the specificity of the episode narrows.
- A Committee member recommendation that the continuum be related to prior NQF work. It must be determined whether the continuum is in sync with the previous NQF Efficiency /Episodes Steering Committee, and, if not, the language must be made similar.
- Some Committee members believed that the Call for Measures should focus on per episode because that is where the biggest gap in measurement is right now.
- The Committee also discussed whether the Call for Measures should include all of these resource use measurement approaches, or focus more narrowly on a few.

Resource inputs in the real world

- A Committee member suggested that a section be added to the white paper that explains that there are inputs in the system that are not currently accounted for because there are no codes for many of them. For other inputs, codes do exist, but they are just not paid for.
- The idea of billing to determine inputs is important, especially when codes are needed to generate bills; but there are some systems that do not generate bills (Kaiser).
- One type of resource input that is frequently an unpaid services is patient education and counseling. This is usually not billable and does not show up in the resource use measures, and therefore goes unaccounted for. Some practices will not offer this service, especially because it is not billable. The use of shared decision tools and evidence-based programs are ways to better account for patient-centered inputs. These should not be excluded as resource inputs.

NATIONAL QUALITY FORUM

- There is often the need to distinguish between resources/inputs “used,” “billed,” and “paid for.” As a best practice for determining actual inputs, reimbursement schedules should not be used to determine costs.
- The Committee questioned whether benefit design should be considered in measuring inputs. Although it makes the discussion very complex, there is evidence that it impacts behavior and use of resources. Is it hopeless to attempt to adjust for this? The Committee agreed that there will be issues that cannot be adjusted for, but they should be addressed in the white paper along with discussion around when risk-adjustment vs. stratification should be used.

Section 3: Types of resource use measures

The Committee was tasked in this section with determining which of these steps would be subject to evaluation to be reflected in the criteria, to identify any missing steps, and to suggest any changes to those already listed.

Measure specification steps

- Data preparation
 - Inclusions and exclusions
 - Data cleaning
- Creation of units for measurement
 - Defining unit for measurement
 - Define clinical or temporal logic
 - Define unit (episode) logic
 - Integrate risk-adjustment (optional)
- Applying units measurement
 - Resource use units specified
 - Costing methodology
 - Estimating resource use or cost
 - Peer group selection
 - Threshold or outlier decision
 - Sample sizes—minimum number units form measurement for feedback or public reporting
 - Calculating comparisons or benchmarks
 - Calculating score results (observed to expected)
 - Attribution of results
 - Risk-adjustment
 - Reporting through descriptive statistics (e.g., distribution, confidence intervals)
 - Setting thresholds—identification and decision around outlier costs of entire episode or patient (high and low)
 - Integrating with quality of care measurement
- Terminology for this process is important. Rather than having three buckets of “steps,” the Committee agreed to describe them as steps for specifying the measures and call the entire process “measure specification,” removing the three buckets around them.
- Developers should be encouraged to consider each of the steps and explain why they choose to apply (or not apply) a step.
- Attribution is a very difficult step, and the white paper somewhat dismisses it as easy. This section should also be expanded to provide a more in-depth discussion.
- There is a need to determine where specifications end and implementation guidance begins.

NATIONAL QUALITY FORUM

Risk-adjustment

- The paper does not provide sufficient discussion of risk-adjustment. Risk-adjustment is very complex, and there are a lot of data that provide a much more in depth discussion of the issues. The white paper should reference this work and highlight where there have been significant advancements in current risk-adjustment methodologies. The white paper should also attempt to reflect current state of risk-adjustment methodologies.
- The Committee suggested that risk should be based on the patient's total universe of claims and not just the claims of the entity being analyzed.
- There should be a basis for data selection, and rationale for those decisions will be needed.

Section 4: Limitations, implications, and open questions

- These resource use measures will not be eligible for time-limited endorsement status. NQF policy does not allow for time-limited status for complex measures such as these.
- Upon submission, measure developers should be required to explain geographic usage and statistical information in geographic regions.
- The white paper should provide specific examples of resource use scores.

DAY 2: TUESDAY, JULY 13, 2010

Meeting Objectives

- Determine the criteria by which resource use measures should be evaluated
- Obtain Committee input on the Call for Measures and measure developer guidance

Resource Use Measure Evaluation Criteria and Principles

Ms. Turbyville started the day with a description of NQF's four main evaluation criteria (importance, scientific acceptability, usability, and feasibility). The Committee was tasked with reviewing the proposed sub-criteria for resource use measures and making any suggestions for changes, expansion, additions, etc.

During a brief Committee discussion on the types of measures that should be included in the Call for Measures, the Committee agreed that the desired level of analysis (health plan provider, regional) for submitted measures should be explicit in the call.

Prior to reviewing each of the individual evaluation criteria, the Committee reviewed each of the resource use evaluation principles that provide the framework and guide the evaluation of resource use measures, and made recommendations for changes and rewording. The purpose of the principles is to provide developers the context for which the evaluation criteria will be specified.

Principles

Principle 1

Resource use measures are measures of input (usually in terms of dollars, but can also be in terms of relative value units); they are not measures of quality. Resource use measures are an important building block to measures of efficiency of care; future measurement efforts should integrate and explicitly incorporate quality or appropriateness performance.

NATIONAL QUALITY FORUM

- The Committee discussed the need to develop resource use measures to arrive at efficiency measures. The Committee suggested that efficiency is a function of quality and resource use and provided the following equation: Efficiency = F(Q,RU)
- Resource use measures are important building block to measures of efficiency of care; “future” measurement efforts should integrate and explicitly incorporate quality or appropriateness of performance.
- Resource use measures are proxies for quality. The best measurement efforts count resource inputs in relation to outcomes.

Principle 2

Efficiency is one of the IOM five quality domains; it is a measure of inputs (amount of resource used per population, or episode, or procedure) per output (quality).

Principle 3

The justification and intended purpose for resource use measures is to examine, and ultimately understand variation and poor performance¹— in regard to the cost of healthcare in the context of quality.

- In dealing with resource use measures the Committee thought it best to eliminate the term “poor performance,” because it implies quality, and instead use the word “practice.”
- The Committee agreed that there is a need to explain variation because there are different types of variation (e.g., geographic).
- Implicit in use is where the money is being spent.
- Measure developers should explain how they came up with their costing, attention to coding practices, etc., and should indicate who the intended users of the measure are.
- The Committee grappled with the language of this principle, considering “...to examine, understand and ultimately reduce unnecessary healthcare costs.”

Principle 4

There is a continuum of types of resource use measures; all types of resource use measures must meet evaluation criteria. (See Continuum Model on page 3.)

- The continuum model should be defined in this principle or reference its discussion somewhere.

Principle 5

The resource use measure calculation must be explicitly stated and transparent such that the approach can be deconstructed and implemented in a standard manner.

- Measures should be comparable across services and implicit in use.

Principle 6

Resource use measure development should begin with a conceptual construct of cost with a set of components (service categories) that one wishes to summarize into one score. The methods used to develop and test the components must be justified.

¹ Poor performance for resource use, in addition to substantial variation, may be defined to include: high cost, rapidly increasing cost, high volume, etc.

NATIONAL QUALITY FORUM

- Remove the term “cost” when referring to constructs. This idea may already be imbedded in the criteria. The set of components refers to service categories.
- Expand the language to explain how a measure that uses a roll up of one score should be defined and justified.
- Consider removing this principle altogether.

Principle 7

Combining multiple service categories into one resource use estimate increases complexity; using methodologically sound methods is of paramount importance. The approach should be fully transparent. Furthermore, even though the background calculations may be more complex, the final resource use score or result should be simple and readily interpretable by all stakeholders.

- In formulating the principles, the Committee considered taking the broadest direction when possible so that comprehensive measures are produced.
- Comprehensive measures are preferable.

Principle 8

Methods for combining the component scores influence the interpretation of the composite measure results and must be justified (e.g., all-or-none scoring indicates whether patients receive all/less than all of the items measured; averaging across component scores may obscure low or high scores of individual components).

- Remove the wording “all or none scoring indicates whether patients receive all/less than all of the items measured.”
- This principle should be simplified and readily interpretable by all stakeholders.
- This principle is already implicit in the existing NQF process for evaluating measures.

Suggested new principles

- Add principle about standardization and comparability across services.
- Measures should address broad populations as much as possible.
- The purpose of this process is to endorse resource use measures, so that better measures are created for understanding efficiency.
- The parts of the measure submitted need to support the overall measurement concept.

Evaluation Criteria

Importance to Measure and Report

- The unit of public reporting must be determined.
- Hospitals could utilize public reporting across the continuum of care.
- Allow for innovation and then decide on public reporting at a later date.
- Indicate if the measures will be used for public reporting.
- Smaller vendors need to be accommodated in order to encourage submission of measures to this project.

NATIONAL QUALITY FORUM

Scientific Acceptability

- 2a: Include the list of resource units/scope of resources being measured.
- 2a: There is a preference for measures that more comprehensively include resources, and should be explicitly indicated by developer as well.
- The measurement framework report should promote shared responsibility and should harmonize measures across settings.
- The Committee was tasked with determining which specification steps (as shown above) should be subject to evaluation and included in the criteria.
- The Committee agreed that it would be more useful to convene a smaller sub-group to discuss these steps in detail and make a recommendation for the full Committee to consider. Seven Committee members volunteered. NQF staff will follow up after the meeting to convene the workgroup for further discussion of these steps.

Feasibility

- 4b: Be explicit about type of electronic source or whether the measure developer needs to use a standard code set, etc.
- 4a and 4b: As written these sub-criteria seem to be limited to claims data. Rewrite to clarify if it should be more inclusive. State, “Electronic data is preferable.” Ask, “What does electronic source include?”

Usability

- It was suggested that measure developers scan for related endorsed quality measures that could be used with the resource use measures. It doesn’t necessarily need to be a requirement within the criteria

Best in Class

- Best in class is determined on the quality side by first evaluating each measure individually against each criteria. If each measure is determined to have met all the criteria, the Committee must then determine which is “best in class” by performing a side-by-side comparison and deciding which is best.
- Some Committee members expressed concern about having to select a “best” measure given the complexity of these measures and their submission packets. What about endorsing individual measures with groupers from different vendors? This will be a challenge.
- Some assessment will be needed for measures focusing on similar areas.
- NQF may identify a best in class and this may be facilitated by the work of the Testing Task Force.
- The Committee would like to leave this criteria open for further discussion and perhaps change the language of this criteria so that it is not required. This is the first effort at endorsing resource use measures, and the field may not yet be ready for a “best in class” during this process.

Review of Call for Measures

Ms. Wilbon led the Steering Committee members through the Call for Measures draft document and asked for feedback regarding any necessary changes and/or additions.

NATIONAL QUALITY FORUM

Public Comment

- A comment from a public audience member stated that the language used in the white paper and evaluation criteria should be more inclusive, that is, they should include nurse practitioners and other clinicians, and not just physicians.
- It was commented that the term “gaming” should not be used in the white paper.

Next Steps

- The White Paper Public and Member Comment Period is tentatively set to open in late August.
- NQF staff will contact the volunteers and convene the workgroup to determine the measure specification steps that should be required and or optional (if any) in the submission of resource use measures. They will also be asked to determine how these measure specifications steps should be reflected in the measure evaluation criteria.
- The next scheduled conference call will be October 5, 2010, at 11 a.m. ET to discuss comments on the white paper.