Health Insurance Marketplace

QRS Methodology: Summary of Key Themes from Public Comment and Changes to Methodology

December 2014

Background

Based on section 1311(c) of the Affordable Care Act,[1] CMS developed the Quality Rating System (QRS) to: inform consumer selection of Qualified Health Plans (QHPs) offered through a Health Insurance Marketplace (Marketplace); facilitate regulatory oversight of QHPs; and provide actionable information to QHPs for performance improvement. CMS also developed the Qualified Health Plan Enrollee Experience Survey (QHP Enrollee Survey), which will yield enrollee experience response data, a subset of which will be used in the QRS.

On March 28, 2014, CMS posted the *Quality Rating System Scoring Specifications* for public comment. The document outlined the draft methodology CMS will use to derive quality ratings for QHP issuers' Marketplace products based on QRS clinical measure and QHP Enrollee Survey response data. The public comment period ended on April 28, 2014. CMS incorporated public comment feedback into the methodology as part of ongoing refinements in preparation for the 2015 QRS beta test. The revised methodology that CMS intends to use during the 2015 QRS beta test is outlined in the *2015 Quality Rating System and Qualified Health Plan Enrollee Experience Survey Technical Guidance* document released in September 2014.[2] The *2015 Technical Guidance* provides information regarding the 2015 beta test of the QRS and QHP Enrollee Survey for QHP issuers.[3]

This document, *QRS Methodology: Summary of Key Themes from Public Comment and Changes to Methodology*, summarizes key themes from the public comments received on the *QRS Scoring Specifications.* It also includes a summary of the refinements made to the QRS methodology (and the method of communicating it) in response to public comments, input from the QRS Technical Expert Panel (TEP) and additional CMS analyses.

Summary of Changes to QRS Methodology

CMS made two types of refinements to the draft methodology described in *the QRS Scoring Specifications* issued for public comment, which are now incorporated in the *2015 Technical Guidance*: 1) changes to the methodology itself to simplify the overall methodology without sacrificing reliability, and 2) language refinements to improve the transparency and overall comprehension of the methodology.

[1] The Patient Protection and Affordable Care Act of 2010 (Pub. L. 111–148) as amended by the Health Care and Education Reconciliation Act of 2010 (Pub. L. 111–309) (collectively referred to as the Affordable Care Act).

[2] *2015 Quality Rating System and Qualified Health Plan Enrollee Experience Survey Technical Guidance* can be found on the CMS Marketplace Quality Initiatives website at http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Health-Insurance-Marketplace-Quality-Initiatives.

[3] The requirements outlined in the *2015 Quality Rating System and Qualified Health Plan Enrollee Experience Survey Technical Guidance* are based on statute and CMS regulation, including Patient Protection and Affordable Care Act; Exchange and Insurance Market Standards for 2015 and Beyond; Final Rule, 79 FR 30240 at 30352 (May 27, 2014) (45 CFR Parts 144, 146, 147, et al.

CMS CENTERS FOR MEDICARE & MEDICAID SERVICES

1. Methodology Changes:
    - *Weighting and Aggregation*
        - Removed reference to an un-weighted indicator average approach for survey measure scoring. CMS will determine the weighting approach to survey measures during the 2015 QRS beta test.
        - Altered the scoring approach to specified measure indicators (Prenatal and Postpartum Care; Proportion of Days Covered) to treat the indicators as distinct measures (vs. use of indicator weighting).
        - Removed the weighting exception for indicators within the Antidepressant Medication Management measure. CMS will average the indicators to create the measure score.
        - Removed the full-scale rule exception for the Patient Safety domain as it relates to the scoring approach for summary indicators. To determine whether the Clinical Quality Management summary indicator is reportable, CMS will apply the full-scale rule without exception meaning all domain scores must be present to calculate the summary indicator score.
        - Reorganized the hierarchy: Moved both "Well-Child Visits" measures (from Access to Preventive Visits composite to Staying Healthy Child composite) and Care Coordination measure (from Clinical Quality Management summary indicator to Enrollee Experience summary indicator).
    - *Cut Points and Performance Thresholds*
        - Removed reference to established cut points of 25, 50, 75, and 90. CMS will determine the cut points during the 2015 QRS beta test.

2. Language Refinements:
    - *Weighting and Aggregation*
        - Revised the visual representation of the hierarchy to improve overall framing (e.g., added composite labels, changed summary indicator name from Member Experience to Enrollee Experience). These visual updates do not impact the rating methodology.
    - *Minimum Denominator Size and Minimum Enrollment Size*
        - Changed references from minimum sample size to minimum denominator size.
    - *Standardization and Benchmarking*
        - Clarified the meaning of standardization of measure scores.
    - *General*
        - Made minor edits throughout to clarify areas of confusion (e.g., changed references from reporting scores and ratings to calculating scores and ratings).

The revised methodology represents CMS' intended methodology to derive quality ratings for the 2015 QRS beta test. During the 2015 QRS beta test, CMS will use reported data from QHP issuers to inform outcomes on remaining methodological decisions, specifically regarding minimum denominator size, cut points, and the weighting approach associated with the QRS survey measure indicators. Additionally, CMS will conduct confirmatory testing and comparative analyses on the beta test results to understand the results produced by the overall methodology (and associated hierarchy[4]).

[4] The hierarchy associated with the revised methodology is provided in the Appendix.

Details on Public Comments and Summary of Changes

CMS received a total of 17 responses from QHP issuers, associations, and various industry stakeholders who responded during the *QRS Scoring Specifications* public comment period. The following provides an overview of key themes from public comment feedback, along with CMS' response, including any associated changes to the QRS methodology.

Weighting and Aggregation

Summary of Public Comment: A weight is a mathematical function that is used when producing a sum or average to give some elements greater or lesser influence than other elements. Respondents commented on the weighting approach for measures and hierarchy components (e.g., composites, domains, and summary indicators). These components represent levels of scores that ultimately produce the global rating. Scores for a higher-level component are derived from aggregating scores of lower-level components in the hierarchy. Respondents requested additional information regarding CMS' approach to aggregation of scores and component weighting (both explicit and implicit). Implicit weighting refers to the fact that for hierarchy components with fewer measures, each individual measure has more impact on an aggregate score than do measures in components with many measures. Public comment respondents also provided specific recommendations as described below:

- CMS should provide a rationale for weighting the Antidepressant Medication Management measure differently.
- Measure weighting should not be equal. Equal weighting will blur important distinctions in plan and provider performance related to aspects of patient care.
- Measures should be weighted based on how measures are collected (e.g., chart review measures should be weighted more heavily), based on the type of measure (e.g., outcome measures should be weighted more heavily than process measures), and/or based on the source of the data (e.g., measures derived from clinical data should be weighted more heavily than measures derived from survey data).
- CMS should abandon the average-of-averages approach and convert the standardized measure scores directly into a final, global, five-star rating.

***Summary of CMS' Response and Changes Made to Refine the Methodology*:** For the revised methodology, CMS will retain the average-of-averages approach. As certain measures will not be reportable until 2016, and other measures may not be reportable due to insufficient denominator size, CMS' choice to use an average-of-averages approach corrects for the differences in availability of measures.

Several QRS measures are comprised of two or more indicators. For QRS survey measures comprised of two or more indicators (or QHP Enrollee Survey questions in this case), CMS will determine the weighting approach during the 2015 QRS beta test. For QRS clinical measures, CMS will use a weighted average method to average each measure's individual indicator rates and calculate a measure score. The "weights" are based on the respective indicator denominator sizes. CMS created an exception for two measures, Prenatal and Postpartum Care and Proportion of Days Covered, whose indicators are treated as unique measures (in that they are weighted equally alongside other measures to form composites). Creating these exceptions does not significantly diminish the reliability of the composite score and clearly shows how the composite is a function of measures that cover different aspects of health.

Additionally, to improve transparency of the overall methodology, CMS removed two exceptions to aggregation rules. First, CMS removed the "down-weighting approach" for the Antidepressant Medication Management measure indicators. Second, CMS removed the full-scale rule exception for the Patient Safety domain as it relates to the scoring approach for summary indicators. These changes removed complexity from the QRS scoring process, while not significantly decreasing measurement reliability.

CMS also made changes to the organization of measures within the QRS hierarchy based on feedback and additional CMS analysis:

- CMS moved both "Well-Child Visits" measures from the Access to Preventive Visits composite to the Staying Healthy Child composite. Stakeholder feedback and CMS analysis found that these measures have more in common with the measures in Staying Healthy Child.
- CMS moved the Care Coordination measure from the Clinical Quality Management summary indicator to the Enrollee Experience summary indicator. The Care Coordination measure assesses an enrollee's experience of coordination and is distinct from the other clinical measures in this summary indicator. QRS TEP members noted that this would also reflect the experience of other performance management programs, which have over the years increasingly grouped together survey measures that assess a consumer's assessment of care.

Minimum Denominator Size and Minimum Enrollment Size

Summary of Public Comment: A minimum sample size is the minimum number of observations needed in order to create a reliable score. As the draft *QRS Scoring Specifications* document did not provide specific guidance on minimum sample sizes, many public comment respondents requested additional details. Respondents gave specific suggestions related to minimum sample size, as listed below.

- Survey sample sizes should not be less than 1,000, as small samples may incorrectly represent the plan's actual performance due to adverse selection or random sampling variation.
- CMS should investigate (via modeling) the fairness of half-scale and full-scale rules for small plans that may not have sufficient data required to receive a global rating (i.e., will have too many non-reportable composites, domains, and/or summary indicators).
- CMS should not require that plans meet continuous enrollment criteria for measures to be scored.
- Reporting and scoring should be at the product level, not metal level.

Summary of CMS' Response and Changes Made to Refine the Methodology: For the revised methodology, CMS will require QHP issuers to submit measure data regardless of denominator size. Measures with an insufficient denominator size will be excluded from QRS scoring. QHP issuers that do not meet the minimum denominator size requirement for a measure will not receive a score for that measure. CMS will establish the minimum denominator size in 2015, when beta test data are available, and will publish these details in future technical guidance. Additionally, CMS is retaining the continuous enrollment criteria as they are set by measure stewards to support reliability and validity of the measurement.

In terms of enrollment size, for both the QRS and QHP Enrollee Survey, QHP issuers are required to collect and submit data for those QHPs that have more than 500 enrollees as of July 1, 2014, in a given product type.[5] For the 2015 beta test, CMS requires data collection and submission by QHP issuer product type (e.g., HMO, PPO). For both QRS clinical measures and survey measures, CMS will explore data collection at a more granular level of QHP issuer coverage in the future, keeping in mind the need to balance the value of this information for consumers with QHP issuer data collection, validation, and reporting efforts.

Standardization and Benchmarking

Summary of Public Comment: Standardization helps consumers compare the QRS ratings for QHPs using a uniform framework. As described below, respondents made several recommendations for how CMS should determine the appropriate reference group for standardization.

[5] The QHP Enrollee Survey minimum enrollment requirement aligns with standards set forth in 45 C.F.R. § 156.1125(b)(1). CMS established the minimum enrollment requirement for QRS to align with QHP Enrollee Survey minimum enrollment requirement and to contribute to a sufficient size for credible and reliable results.

- A national reference group should not be used for benchmarking. Regional or local benchmarks are preferred due to consumer interest and relevancy. Both national and regional benchmarks could be calculated and displayed.
- A phase-in solution should be used, with an initial approach that combines national and regional scores and over time moves to a score based solely on a national reference group.
- Standardization does not contribute any information regarding improvement and masks whether the highest-rated plans are "good" only relative to their peers (but are still poor clinically) and whether the lowest-rated plans are "bad" only relative to their peers (but are high clinical performers nonetheless).
- CMS should provide additional information to explain how reference group data is handled over multiple years. Specifically, CMS should clarify whether reference group data from one or multiple years would be used, whether re-standardization would occur each year, and whether standards will remain fixed to help show improvement.
- Using percentiles, rather than absolute values, is problematic, because it misrepresents perceived performance when QHPs are grouped at the upper end of the performance range. Absolute values will ensure that there are statistically significant and meaningful differences among the five-star ratings, reflecting true differences in plans' performance.
- Using percentiles, rather than absolute values, is preferred as this approach gives consumers meaningful comparisons both within their markets and to other markets across the country. Additionally, using percentiles also provides for normal performance distributions on measures that will reflect both current measure performance and performance improvement.
- Before final adoption, CMS should test the effects of combining product types to create the national peer group.

Summary of CMS' Response and Language Refinements to the Methodology: For the revised methodology, CMS will retain the national reference group, meaning QRS percentile ranks will be based on one national, all-product reference group. A national reference group is consistent with CMS policy and performance measurement programs (e.g., Medicare Stars), and is likely the most statistically robust choice in the initial years. CMS will standardize measure scores by calculating national percentile ranks before calculating composite and higher-level QRS component scores. For each measure with a reportable rate, CMS will use the calculable QHP product's rate to create national percentile ranks.

Cut Points and Performance Thresholds

Summary of Public Comment: A cut point is a numeric score value that serves as a threshold to delineate a category or level of performance. These levels of performance produce the star rating scale. Respondents expressed concern that the proposed cut points (25/50/75/90) may not be optimal, and made specific recommendations as described below.

- The proposed cut point may not enable enough plans to receive high ratings (e.g., five stars on the proposed five-star scale). CMS should revise the cut points based on statistical analysis and empirical data and distribute more evenly (e.g., 20/40/60/80 or 25/50/75).
- CMS should assess through sensitivity analysis the effects of different cut points.
- CMS should consider using half-star designations.

Summary of CMS' Response and Changes Made to Refine the Methodology: CMS intends to retain the five-star rating system with no half stars during the beta test to mitigate against the risk of misclassification (which typically increases with additional categories/stars). In addition, the use of a five-star rating system aligns with existing CMS rating programs (e.g., Medicare Stars, Nursing Home Compare). CMS will determine the placement of cut points during the 2015 QRS beta test. Using beta test data, CMS will test potential approaches to assess the sensitivity of the results to different cut points.

Appendix: QRS Hierarchy for Revised QRS Methodology

The table below illustrates the QRS hierarchy, which is the organization of measures into composites, domains, and summary indicators, and which contributes to a single global rating.

QRS Summary Indicator	QRS Domain	QRS Composite	Measure Title	NQF #
Clinical Quality Management	Clinical Effective-ness	Asthma Care	Medication Management for People With Asthma (75% of Treatment Period)	1799
		Behavioral Health	Antidepressant Medication Management	0105
			Follow-Up After Hospitalization for Mental Illness (7-Day Follow-Up)	0576
			Follow-Up Care for Children Prescribed ADHD Medication	0108
			Initiation and Engagement of Alcohol and Other Drug Dependence Treatment	0004
		Cardiovascular Care	Controlling High Blood Pressure	0018
			Proportion of Days Covered (RAS Antagonists)	0541
			Proportion of Days Covered (Statins)	0541
		Diabetes Care	Comprehensive Diabetes Care: Eye Exam (Retinal) Performed	0055
			Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Control (<8.0%)	0575
			Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Testing	0057
			Comprehensive Diabetes Care: Medical Attention for Nephropathy	0062
			Proportion of Days Covered (Diabetes All Class)	0541
	Patient Safety	Patient Safety	Annual Monitoring for Patients on Persistent Medications	Not Endorsed*
			Plan All-Cause Readmissions	1768
	Prevention	Checking for Cancer	Breast Cancer Screening	Not Endorsed*
			Cervical Cancer Screening	0032
			Colorectal Cancer Screening	0034
		Maternal Health	Prenatal and Postpartum Care (Postpartum Care)	1517
			Prenatal and Postpartum Care (Timeliness of Prenatal Care)	1517
		Staying Healthy Adult	Adult BMI Assessment	Not Endorsed
			Chlamydia Screening in Women	0033
			Aspirin Use and Discussion	Not Endorsed
			Flu Vaccinations for Adults Ages 18-84	0039
			Medical Assistance With Smoking and Tobacco Use Cessation	0027
		Staying Healthy Child	Annual Dental Visit	1388
			Childhood Immunization Status (Combination 3)	0038
			Human Papillomavirus Vaccination for Female Adolescents	1959
			Immunizations for Adolescents (Combination 1)	1407
			Weight Assessment and Counseling for Nutrition and Physical Activity for Children and Adolescents	0024
			Well-Child Visits in the First 15 Months of Life (6 or More Visits)	1392
			Well-Child Visits in the Third, Fourth, Fifth, and Sixth Years of Life	1516

* Measure not NQF endorsed as of September 2014, but was submitted in early 2014 for endorsement.

QRS Summary Indicator	QRS Domain	QRS Composite	Measure Title	NQF #
Enrollee Experience	Access	Access to Care	Access to Care	0006**
	Care Coordination	Care Coordination	Care Coordination	Not Endorsed
	Doctor and Care	Doctor and Care	Cultural Competence	Not Endorsed
			Rating of All Health Care	0006**
			Rating of Personal Doctor	0006**
			Rating of Specialist	0006**
Plan Efficiency, Affordability, & Management	Efficiency & Affordability	Efficient Care	Appropriate Testing for Children With Pharyngitis	0002
			Appropriate Treatment for Children With Upper Respiratory Infection	0069
			Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis	0058
			Use of Imaging Studies for Low Back Pain	0052
	Plan Service	Enrollee Experience with Health Plan	Access to Information	0006**
			Plan Administration	0006**
			Rating of Health Plan	0006**
Collected but Not Included for Purposes of QRS Scoring and Ratings				
N/A	N/A	N/A	Relative Resource Use for People with Diabetes (Inpatient Facility)	1557

**NQF ID #0006 reflects NQF endorsement for the CAHPS® Health Plan 4.0 Survey. The QHP Enrollee Experience Survey and associated QRS survey measures largely align with items from the CAHPS® Health Plan 5.0 Surveys, which were not yet been submitted for endorsement as of September 2014. Further, the Plan Administration survey measure includes one survey item developed by CMS; this survey item is not included in the CAHPS® Survey.