

DEPARTMENT OF HEALTH & HUMAN SERVICES  
Centers for Medicare & Medicaid Services  
7500 Security Boulevard  
Baltimore, Maryland 21244-1850



## **CENTER FOR MEDICARE**

---

**DATE:** November 21, 2014

**TO:** All Medicare Advantage Organizations, Prescription Drug Plan Sponsors, and Other Interested Parties

**FROM:** Amy K. Larrick, Acting Director, Medicare Drug Benefit and C & D Data Group

**SUBJECT:** Request for Comments: Enhancements to the Star Ratings for 2016 and Beyond

Proposed methodology enhancements and changes for the 2016 Star Ratings for Medicare Advantage (MA) and Prescription Drug Plans (PDPs) are described in this document. Advanced notice of potential changes for the 2017 Star Ratings and beyond are also included. MA Organizations, PDP Sponsors, advocates, and other stakeholders have this opportunity to provide comments in advance of the draft 2016 Call Letter. Because the timelines for the annual draft Call Letter, combined with the statutory deadlines for the Advance and Final Rate Notices, do not provide CMS sufficient time to fully explore substantive changes suggested by commenters, this Request for Comments (R4C) allows CMS additional time to review and evaluate comments prior to the Call Letter process.

CMS structured the current Star Ratings strategy to be consistent with the six priorities in the National Quality Strategy. The six priorities include: making care safer by reducing harm caused by the delivery of care; ensuring that each person and family are engaged as partners in their care; promoting effective communication and coordination of care; promoting the most effective prevention and treatment practices for the leading causes of mortality; working with communities to promote wide use of best practices to enable healthy living; and making quality care more affordable for individuals, families, employers, and governments by developing and spreading new health care delivery models. The measures span five broad categories, including:

- Outcome measures that focus on improvement to a beneficiary's health as a result of care that is provided;
- Intermediate outcome measures that concentrate on ways to help beneficiaries move closer to achieving a true outcome;
- Patient experience measures that represent beneficiaries' perspectives about the care they receive;

- Access measures that reflect processes or structures that may create barriers to receiving needed health care; and
- Process measures that capture a method by which health care is provided.

The Star Ratings help inform beneficiaries about the performance of health and drug plans on the Medicare Plan Finder website, as well as serve as the basis of Quality Bonus Payments (QBPs) for MA organizations. CMS continues to improve the Part C and Part D quality and performance measurement system to focus on beneficiary outcomes, beneficiary satisfaction, population health, and health care efficiency. The goal is that the Star Ratings system will not only influence beneficiaries' plan choices but also drive organizations and sponsors toward higher quality and more efficient care.

Star Ratings is a year-round process for both CMS and sponsors.

- CMS provides guidance on methodology changes anticipated for the 2016 Star Ratings and advanced notice of potential changes for the 2017 Star Ratings and beyond through this fall's R4C and then in the draft 2016 Call Letter in February 2015.
- Sponsors and other stakeholders should actively participate by submitting comments regarding the potential changes to the Star Ratings methodology or submit additional comments to propose other changes not mentioned in the R4C. Comments to the R4C inform proposals in the draft Call Letter, and comments to the draft Call Letter inform the final Call Letter.
- After consideration of all comments, CMS announces the methodology for the 2016 Star Ratings in the final 2016 Call Letter in April 2015.
- Outside of the Call Letter process, sponsors should review the underlying data used for the individual Star Ratings measures as available throughout the year, and notify CMS of any errors or questions in a timely manner.
- Each summer, CMS holds a Part C and D User call with plans to provide updates to the upcoming Star Ratings release.
- CMS conducts two plan preview periods via HPMS to identify any necessary data corrections or revisions to our draft Technical Notes. During these preview periods, CMS expects sponsors to raise concerns about their raw measure data and Star Ratings. Changes to the methodologies for measure calculations or Star Rating calculations cannot be made during this time.
- In November, CMS offers MA organizations an appeals process for QBPs. The administrative review process is a two-step process that includes a request for reconsideration and a request for an informal hearing after CMS has sent the MA organization the reconsideration decision. This process may only be initiated on the basis of a calculation error (miscalculation) or a data inaccuracy (incorrect data).

Comments to this R4C should be submitted via the following link:

<https://cmsgov.wufoo.com/forms/enhancements-to-the-star-ratings-2016-and-beyond/>.

Please submit only one set of responses per organization. If your organization has comments on more sections than the form allows, you may submit the form more than once. If you wish to submit additional supporting documents, you may send via email to:

[PartCandDStarRatings@cms.hhs.gov](mailto:PartCandDStarRatings@cms.hhs.gov).

Comments submitted by Wednesday, December 17<sup>th</sup> at 5pm ET will be considered as we finalize proposed changes for the 2016 Star Ratings for the draft 2016 Call Letter. Stakeholders will have another opportunity to comment on the 2016 Star Ratings methodology and proposed changes through the draft Call Letter process. The same comments should not be resubmitted to the draft Call Letter.

Questions related to this R4C may be sent to: [PartCandDStarRatings@cms.hhs.gov](mailto:PartCandDStarRatings@cms.hhs.gov).

Thank you for your participation.

## Proposed Enhancements to the 2016 Star Ratings and Beyond

One of CMS' most important strategic goals is to improve the quality of care and general health status of Medicare beneficiaries. For the 2016 Star Ratings, CMS continues to enhance the current methodology so it further aligns with our policy goals. Our priorities include enhancing the measures and methodology to more reflect the true performance of organizations and sponsors, maintaining stability due to the link to payment, and providing advance notice of future changes. In this document, we describe the enhancements being considered for the 2016 Star Ratings and beyond. Unless noted below, we do not anticipate methodology changing from the 2015 Star Ratings.

For reference, the list of measures and methodology included in the 2015 Star Ratings is described in the Technical Notes: <http://go.cms.gov/partcanddstarratings>.

The star cut points for all measures and case-mix coefficients for the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey and Health Outcomes Survey (HOS) will be updated for 2016 with the most current data available.

As announced in previous years, we will annually review data quality across all measures, variation among organizations and sponsors, and measures' accuracy and validity before making a final determination about inclusion of measures in the Star Ratings.

CMS is exploring the development of an integrated Star Rating system for Medicare-Medicaid Plans (MMPs) participating in the capitated financial alignment model. The purpose of this effort is to develop a rating system that acknowledges the additional needs of Medicare-Medicaid enrollees and measure the performance of the MMPs in integrating the Medicare and Medicaid benefits. More information will be provided during the first quarter of 2015.

### **A. Changes to the Calculation of the Overall Rating and the Part C and D Summary Ratings**

#### **a. Background**

CMS is interested in improving the accuracy of the assignment of overall and Part C and D summary Star Ratings and ensuring the system creates incentives for quality improvement. In constructing Star Ratings, a key concern is the potential for generating Star Ratings that do not reflect a contract's "true" performance, otherwise referred to as the risk of "misclassification." For example, a contract's performance (e.g., scoring a "true" 4-star contract as a 3-star contract, or vice versa). Misclassification occurs in any measurement system because all performance measurement is a mixture of *signal* (true performance) and *noise* (random measurement error due to rounding, sampling variation and similar factors). Over the years several features have been implemented in the quality rating system to simplify information for consumers, as well as to make the information more transparent for organizations and sponsors. For example, we group the measure scores into star categories and round the data to make it easier for consumers to understand what a particular score means. Since the 2011 Star Ratings, we have also implemented pre-determined 4-star thresholds for some measures to increase transparency for

organizations and sponsors and set expectations for high performance. However, all of these features create more “noise” or measurement error in the system.

## **b. Current Scoring Method**

The 2015 overall Star Rating is a composite measure constructed from 33 measures for Part C and 13 measures for Part D. The measures are numeric scores such as counts and percentages of screening and testing, chronic care, patient experience, customer service, and patient safety measures. Currently, each measure is assigned a rating from 1-5 stars. Scores are grouped using statistical techniques to minimize the distance between scores within a grouping (or “cluster”) and to maximize the distance between scores in different groupings.

There are two methods for calculating the measure stars:

- Clustering. Clusters are defined as contracts with similar distances between their data values and the center data value. The measure scores are inputs for a clustering algorithm, which determines break points in the distribution and groups the scores into star categories.
- Significance testing. The measure scores are assigned stars with a combination of percentile-based categories and whether the score is significantly different from the mean of all contracts.

For the 2015 Star Ratings, 22 Part C and 5 Part D measures have pre-determined 4-star thresholds (67% of Part C measures, and 39% of Part D measures). We did not introduce any new 4-star thresholds for the 2015 Star Ratings. For those measures with pre-determined 4-star thresholds, any contract with a measure score above the threshold receives 4 or 5 stars, and any contract with a score below the threshold receives 1, 2, or 3 stars. The pre-determined 4-star threshold is applied before the clustering or significance testing. For example, for clustered measures, first the contracts that score above the pre-determined threshold are selected, and then this subset is clustered into two categories to determine which contracts receive 4 stars and which receive 5 stars.

Performance consistency across measures is considered an important indicator for the reliability of quality measurement. The individual measures selected by CMS for Star Ratings are proxies for the underlying central concept of high quality care. As such, consistently high performance across all measures is an indication that we can be more confident that an organization or sponsor’s underlying operations and clinical services reflect the high quality of care they provide. In contrast, an organization or sponsor that demonstrates more erratic behavior in measures may not offer the same consistent quality, due to non-aligned operations or clinical services. An organization or sponsor’s inconsistent performance—high on some measures, low on others—could also mean mismanagement of some areas by internal staff or subcontractors.

To incorporate this consistency indicator into the rating process, CMS has applied a “Reward Factor”, previously called an i-Factor, to the mean overall and Part C and D summary ratings since 2009 in order to reward contracts if they have both consistently

high and stable relative performance. Specifically, the Reward Factor calculation adds a value of 0, 0.1, 0.2, 0.3, or 0.4 to each contract's overall and summary ratings according to the variability and mean performance of its measure stars, and in doing so it increases the number of contracts at the high end of the rating scale that have low variation and high mean performance in their individual measure scores. The 2015 Part C & D Star Rating Technical Notes provides more information about the calculations.

### c. Pre-determined Thresholds

Some sponsors and stakeholders are concerned that it is difficult to improve without published targets for achieving 4 or more stars on a measure. While we understand the perceptions that pre-determined 4-star thresholds provide stability by setting performance expectations, in reality the use of pre-determined thresholds violates our principle of assigning stars that maximize the difference between star categories. Pre-determined 4-star thresholds can thus cause contracts to receive different ratings when there is no significant difference in their scores (e.g., if a 4-star threshold is 80%, a contract that scores 79.4% would receive 3 stars while a contract that scores 80.1% would receive 4 stars when there may be no meaningful difference between a score of 79.4 and a score of 80.1). The use of pre-determined 4-star thresholds is also problematic when there is general improvement in measure performance over time or when there are changes to a measure's specifications. It is also problematic when there are large distributional changes in the scores across contracts. In this case, there may not be any contracts with 4 or 5 stars, or any contracts with 1, 2, or 3 stars, for a particular measure. These examples illustrate how pre-determined thresholds increase noise in the Star Ratings and are counter to industry feedback that thresholds assist sponsors in targeting their improvement efforts.

CMS' analyses of past Star Ratings found that sponsors on average have more significant levels of improvements in Part C and D measures **without** pre-determined thresholds, as compared to measures where there are pre-set thresholds. Using the 2015 Star Ratings, our analysis showed that on average only 28% of contracts improved significantly across the 20 Part C measures with 4-star thresholds included in the improvement measure, compared to 51% of contracts that improved significantly across the nine Part C measures without 4-star thresholds. We found similar findings for Part D, where on average, only 24% of contracts showed significant improvement across the five measures with 4-star thresholds included in the improvement measure, while 63% of contracts showed significant improvement across the five Part D measures without 4-star thresholds. These findings continue to show pre-set thresholds hamper continuous quality improvement in MA and Part D contracts.

As announced in the 2015 Call Letter, based on CMS' analyses, we propose removing the pre-determined measure thresholds for the 2016 Star Ratings. The cut points would be determined using the same methodology used in the past (e.g., relative distribution and clustering of the data), and we would continue to use the "Reward Factor" for contracts with consistently high performance.

We understand that some sponsors are concerned that eliminating pre-determined 4-star thresholds will make it more difficult to set targets for performance. Currently, 33% of the Part C measures and 61% of the Part D measures do not have pre-determined 4-star thresholds. As described earlier, 2015 Stars demonstrated as in past years that sponsors achieve more significant improvements in measures without pre-determined 4-star thresholds. We welcome input as to whether we should phase-out the elimination of pre-determined thresholds. For example, in 2016, we could eliminate the pre-determined 4-star thresholds for process measures and for 2017, the remainder of the measures, versus eliminating all pre-determined 4-star thresholds in 2016.

CMS has also considered alternatives to address industry's request for stability via pre-announced benchmarks. One alternative would be to modify the current pre-determined 4-star thresholds by adding an annual improvement percentage increase (IPI) that reflects the trends in improvement we have seen from the 2014 to 2015 Star Ratings. If the national average score for a measure did not improve from the 2014 to 2015 Star Ratings, the current pre-determined 4-star threshold would remain for the measure. In an effort to maintain stability, the IPIs would not be adjusted more than once every three years. Although this alternative would provide stability and further incentivize contracts to improve, it would not solve the problem of misclassification of contract scores in the Star Ratings system due to pre-determined thresholds.

For example, for the Colorectal Cancer Screening measure we have seen the national average increase 4.8% from the 2014 to 2015 Star Ratings. The current 4-star threshold for this measure is 58%. In this alternative we would increase the 4-star threshold to 61% for the 2016 Star Ratings, 64% for the 2017 Star Ratings and so on.

2015 National Average = 65%

2014 National Average = 62%

$IPI = ((2015 \text{ National Average} - 2014 \text{ National Average}) / 2014 \text{ National Average}) * 100$   
 $((65 - 62) / 62) * 100 = 4.8\%$

2016 4 Star Threshold = 2015 4 Star Threshold \* (1 + IPI)

2016 4 Star Threshold = 58 \* 1.048 = 61%

2017 4 Star Threshold = 2016 4 Star Threshold \* (1 + IPI)

2017 4 Star Threshold = 61 \* (1 + IPI) = 64%

We welcome comments on this alternative approach.

In 2014, we provided contract-specific information on the impact of removing pre-determined 4-star thresholds, as well as results of our analyses of performance trends in Star Rating measures, and as applicable, pre-determined 4-star thresholds. In January 2015, through HPMS, we will again provide contracts with these simulations using the 2015 Star Ratings data. A document showing trends overtime in cut points is available at <http://go.cms.gov/partcanddstarratings>. We will continue to update this document to help sponsors target their quality improvement efforts.

**B. New 2016 Measure:**

CMS intends to add the following measure to the 2016 Star Ratings.

1. *Medication Therapy Management Program Completion Rate for Comprehensive Medication Reviews (Part D)*. This measure is based on the PQA-endorsed measure, Completion Rate for Comprehensive Medication Review (CMR), which measures the percentage of beneficiaries who met eligibility criteria for the Medication Therapy Management (MTM) program and who received a CMR with a written summary in CMS standardized format. Since this is a first year measure, it will be assigned a weight of “1”; in future years it will continue to receive a weight of “1” as a process measure. The specifications from the 2015 Display Measure will continue to be used for the 2016 Star Rating. The denominator is the number of beneficiaries who were at least 18 years or older as of the beginning of the reporting period and who were enrolled in the MTM program for at least 60 days during the reporting period. Only those beneficiaries that meet the contracts’ specified targeting criteria per CMS – Part D requirements pursuant to §423.153(d) of the regulations at any time in the reporting period are included in this measure. Beneficiaries that were in hospice at any point during the reporting period are excluded. The numerator is the number of beneficiaries included in the denominator who received a CMR at any time during the reporting period. Sponsors are reminded that they should not restrict their MTM eligibility criteria to limit the number and percent of beneficiaries who qualify for these programs and to whom they must offer a CMR.

**C. Additional 2016 Star Ratings Measures:**

CMS intends to return these measures to the 2016 Star Ratings.

1. *Breast Cancer Screening (Part C)*. The HEDIS specification for 2014 changed the age range from 40 to 69 years old to 50 to 74 years old and increased the numerator time frame for documentation of a mammogram from 24 months to 27 months. These changes were a result of NCQA’s measure re-evaluation process that included: a scan of clinical guidelines and evidence; feedback from variety of stakeholders, including women’s health experts, clinicians, consumer advocates, and health plans; and a public comment period. The revised age range aligns with current recommendations from the U.S. Preventive Services Task Force (Grade B recommendation), American Academy of Family Physicians, and others. The increased numerator time frame from 24 to 27 months provides a 3-month grace period to account for logistics of obtaining a mammogram and is in response to concerns that the lack of a grace period results in women being screened more often than every two years. This change in specifications aligns the measure with the clinical guidelines that were first available in 2009. We are returning this measure to the 2016 Star Ratings, after moving it to the 2015 Display Page for one year since the measure specification changed during the 2013 measurement year expanding the members included in the denominator. Since this is a process measure, it will continue to be assigned a weight of “1”.

2. *Call Center – Foreign Language Interpreter and TTY Availability measures (Part C & D)*. These measures were removed from the 2015 Star Ratings due to concerns about data quality found during the first plan preview. For the 2016 Star Ratings, we plan to collect data for a two month period between November 2014 and June 2015. Data collection will not cross contract years. Since this is an access measure and there is no change in methodology, it will be assigned a weight of “1.5”.
3. *Beneficiary Access and Performance Problems (Part C & D)*. Based on feedback from plans and the changing methodology for calculating the audit results, we are planning to remove audits from this measure to ensure stability in the specifications and include it in the 2016 Star Ratings. The data currently displayed on the 2015 Display Page uses this revised methodology. Appendix A includes the detailed specifications. We will assign this measure a weight of “1” as the methodology change causes this to be a “new” measure for weighting purposes. For the 2017 Star Ratings, it will revert to its weight of “1.5” as an access measure.

#### **D. Changes to Measures for 2016**

CMS’ general policies regarding specification changes to Star Ratings Measures:

- If a specification change to an existing measure is announced in advance of the measurement period, the measure remains in the Star Ratings; it will not be moved to the display page.
- If the change is announced during the measurement period that significantly expands the denominator or population covered by the measure, the measure is moved to the display page for at least one year.
- If the change is announced during the measurement period that does not significantly impact the numerator or denominator of the measure, the measure will continue in the Star Ratings. For example, when during the measurement period additional codes are added that would increase the number of numerator hits for a measure.

The methodology for the following measures is being modified:

1. *Controlling Blood Pressure (Part C)*. In December 2013, the panel members appointed to the eighth Joint National Committee (JNC 8) released updated guidance for the treatment of hypertension. The new recommendations set the treatment goal for patients 60 years of age and older to <150/90 mm Hg and keeps the treatment goal for patients ages 18-59 years at <140/90 mm Hg. The latest guideline also recommends that all diabetic patients age 18 and older should be treated to a goal of <140/90 mm Hg and calls into question the use of other targets.

All HEDIS measures are re-evaluated when there is a change in evidence and/or guidelines. NCQA staff worked with the NCQA advisory committees, including the Cardiovascular Measurement Advisory Panels, Technical Measurement Advisory Panel, and additional stakeholders. The revised measure went to public comment in February-March 2014. The changes were approved by the Committee on Performance

Measurement and Board of Directors in June 2014. The change includes age and condition specific goals to align with JNC 8 guidance: 18-59 years (< 140/90 mm Hg); 60-85 years with diabetes (<140/90 mm Hg); 60-85 years without diabetes (<150/90 mm Hg); and the total rate will be used for reporting and comparison across organizations. We propose to use the updated measure for the 2016 Star Ratings, and this measure will not be transitioned to the Display Page because beneficiaries that meet the old guidelines will automatically meet the newer more lenient guidelines.

2. *Plan Makes Timely Decisions about Appeals (Part C)*. Effective January 2014, organizations are responsible for reviewing dismissal requests and making the decision, rather than forwarding requests to the Independent Review Entity (IRE) for the dismissal decision. Therefore, the IRE no longer captures data around the timeliness of dismissal cases, and consequently, we propose to exclude dismissals from this measure for the 2016 Star Ratings.
3. *Plan All-Cause Readmissions (Part C)*. This is a measure of the percentage of hospital discharges that result in a readmission for any cause within 30 days of discharge. This measure is reported as a ratio of a health plan's observed rate of readmission compared to an expected rate of readmission based on a risk-adjusted model. As discussed in last year's R4C, NCQA has made two changes to this measure which we propose to use for the 2016 Star Ratings: 1) excluding planned readmissions from the measure and 2) removing the current exclusion from the denominator for hospitalizations with a discharge date in the 30 days prior to the Index Admission Date.
4. *Osteoporosis Management in Women who had a Fracture (Part C)*. This measure assesses the percentage of women who had a fracture and received either screening or treatment for osteoporosis. NCQA has added an upper age limit, extended the look back period for exclusions due to prior bone mineral testing, removed estrogens from this measure, and removed single-photon absorptiometry and dual-photon absorptiometry tests from the list of eligible bone-density tests. We propose using the modified measure for the 2016 Star Ratings. For this measure, the denominator changes make the measure easier to meet, while the numerator changes should have very little impact on the measure. Estrogens are in the High Risk Medication in the Elderly measure as drugs to be avoided so they are not commonly being used for treating osteoporosis.
5. *Complaints about the Health/Drug Plan (CTM) (Part C & D)*. CMS proposes to modify the measurement period from 6 months of the current contract year to 12 months of the prior contract year. For example, 12 months of 2014 complaints data will be used for the 2016 measures. Expansion of the data used for this measure will provide a more comprehensive evaluation of the plan. Currently complaints filed in the second half of a year are not accounted for in a contract's performance rating when only the 6-month period is used. Also, this change addresses some contracts' concerns and allows for an approximately 6-month "wash out" period to account for any adjustments per CMS' CTM Standard Operating Procedures.

6. *Improvement measures (Part C & D)*. Please refer to Appendix B for the measures to be used to calculate the 2016 improvement measures.
7. *Appeals Upheld (Part D)*. We propose to modify this measure from using a 6-month snapshot to use a 12-month measurement period as the Part D Appeals Auto-forward measure. For the 2016 Star Ratings measure, instead of using 6 months of 2015 data, we propose using the full 12 months of 2014 data. This change will allow consistency between the Part C and Part D appeals measures as well as expand the measurement period for a more comprehensive evaluation of plans' decisions. Additionally, this change will allow CMS to include reopened cases. Consistent with the Part C measure, if a reopened case is decided prior to April 1 of the following year, the decision for the reopened case is used in place of the reconsideration decision. Previously, contracts with fewer than 5 total cases were not rated in this measure. We will re-evaluate and adjust the minimum number of cases as necessary.
8. *Medication Adherence (for Diabetes Medications and Hypertension (RAS antagonists)) and Diabetes Treatment (Part D)*. PQA updated their 2014 specifications to exclude End-Stage Renal Disease (ESRD) patients from the denominator of these measures based on the ICD-9 code 585.6 and/or by the RxHCC 121. As stated in the 2015 Call Letter, CMS proposes to use the beneficiary ESRD coverage start and termination dates reported in the Medicare Enrollment Database (EDB) rather than the ICD-9 code or RxHCC to identify beneficiaries for exclusion for the 2016 Star Ratings.

EDB data are available for all Part D beneficiaries and, are also current (after considering data lag), whereas RxHCCs do not reflect current diagnoses. CMS' testing of these indicators found a very high level of overlap between the ESRD indicators in the EDB and ICD-9 codes in in-patient and out-patient claims when calculating the final rates for these measures for purposes of the Star Ratings. While there is some lag in data updates, we found the overlap between the two data sources was greater than 95%.

9. *Medication Adherence (Diabetes Medications, Hypertension (RAS antagonists), and for Cholesterol (Statins)) (Part D)*. Currently, when calculating the Proportion of Days Covered (PDC) for the three Adherence measures, if a beneficiary disenrolls from his/her contract in the middle of the calendar year due to death or disenrollment, CMS uses the Common Medicare Environment (CME) enrollment table to obtain the beneficiary's disenrollment date and identify the end of the beneficiary's measurement period. The disenrollment date in the CME is always the last day of the month of disenrollment, regardless of the date of death or actual disenrollment. For example, if a beneficiary is enrolled in a contract starting January 1, 2013 and has a death date of May 10, 2013, CMS uses the May 31, 2013 CME disenrollment date as the end of the beneficiary's measurement period. In response to sponsor feedback, we investigated the feasibility and impact of using the exact death date when available in CME instead of the CME disenrollment date as the end of the beneficiary's measurement period.

This change affects two aspects of the Adherence rate calculation. First, it may reduce the number of beneficiaries eligible for inclusion in the denominator due to the 91 days

restriction. To be included in the denominator of the Adherence rate per the PQA specifications, the beneficiary must have at least two fills of the relevant medication(s) and the first fill must occur at least 91 days before the end of the beneficiary's measurement period. By using the death date instead of the month-end date as the end of the beneficiary's measurement period, some beneficiaries may no longer be eligible for the denominator.

Secondly, for beneficiaries who have death dates that occur before the end of the month, the methodology change shortens the beneficiary measurement period in the PDC calculation. The PDC represents the proportion of days covered by the relevant medication(s) between the date of the beneficiary's first fill and the last day of the measurement period.

Based on simulations with the data used for the 2015 Star Ratings, we found replacing the month-end date with the death date to generally have no effect on the majority of contracts' Adherence rates. This change could have an impact on a small number of individual beneficiaries' PDCs within a contract; therefore, some contracts may observe a small positive or negative impact on their Adherence rates. Simulations of this change using data from the 2015 Star Ratings found that a small number of contracts (less than 5%) may have small increases or decreases in their highest Star Rating. We propose using the exact death date (when available in CME) instead of the CME disenrollment date as the end of the beneficiary's measurement period beginning with the 2016 Star Ratings to improve the specificity of the PDC calculation. We also plan to implement this change in the Patient Safety monthly reports of 2014 PDE in early 2015. We note that there can be up to a three month delay for a beneficiary's death date to populate in the CME; therefore the data may change by the time data are finalized for the 2016 Star Ratings.

10. *Obsolete NDCs (Part D). For the 2016 Star Ratings and display measures (using 2014 PDE data), we propose to implement PQA's 2014 obsolete date methodology.*

Specifically, the obsolete date methodology includes the following steps:

1. Query the MediSpan and First DataBank databases to develop an NDC list.
2. Cross-check the NDC list developed at step 1 against the FDA's Comprehensive NDC Structured Product Labeling (SPL) Data Elements File (NSDE) and its effective dates.
3. Include the NDC in the file if:
  - There is no obsolete date noted by MediSpan or First DataBank or NSDE; or
  - The obsolete date in any of the databases is within the measurement year; or

*The obsolete date is within six months prior to the beginning of the measurement year.*

11. *CAHPS (Part C & D).* As announced in the 2015 Call Letter, we will make minor modifications to the CAHPS methodology to permit low-reliability contracts to receive 5 stars or 1 star. In the past we have not assigned contracts that had a score with low

reliability 1 or 5 stars given the imprecision around the score. However, CMS has conducted additional analyses and some contracts with scores that have low reliability nonetheless have good evidence of performance that is well above the 4-star cut point or below the 2 star cut point. We will modify the CAHPS methodology to permit low-reliability contracts to be assigned 5 stars if the measure score exceeds the 5-star cut point and also exceeds the 4-star cut point by 1 standard error. Similarly, low-reliability contracts can be assigned 1 star if their score is below the 1 star cut point and also falls below the 2-star cut point by 1 standard error.

#### **E. Retirement of Measures**

Due to the release of the new American College of Cardiology (ACC)/American Heart Association (AHA) Guidelines on the Treatment of Blood Cholesterol, NCQA convened its Cardiovascular Measurement Advisory panel in order to address the question of whether changes were needed in their HEDIS measures related to LDL-C control. For HEDIS 2015, NCQA retired the following measures so they will no longer be included in the Star Ratings:

- Cardiovascular Care: Cholesterol Screening
- Diabetes Care: Cholesterol Screening
- Diabetes Care: Cholesterol Controlled

#### **F. Temporary Removal of Measures from Star Ratings**

*Improving Bladder Control (Part C)*. This measure, collected through the Health Outcome Survey (HOS), assesses the percentage of beneficiaries with a urine leakage problem who discussed their problem with their provider and received treatment for the problem. NCQA is making three changes to this measure. First, NCQA changed the denominator of both indicators to include all adults with urinary incontinence, as opposed to limiting the denominator to those who consider urinary incontinence to be an issue. This will remove a potential bias toward only sampling patients who were treated unsuccessfully. Second, NCQA changed the treatment indicator to assess whether treatment was discussed, as opposed to it being received. This will change the measure focus from receiving potentially inappropriate treatments, which often have adverse side effects, to shared decision making between the patient and provider about the appropriateness of treatment. Third, NCQA added an outcome indicator to assess how much urinary incontinence impacts quality of life for beneficiaries. This outcome indicator will not be part of the Star Rating system until additional analyses have been done. These changes required revising the underlying survey questions in HOS. The revised questions will be first collected in 2015. As a result of these changes, there will be no data for this measure for the 2016 and 2017 Star Ratings.

#### **G. Contracts with Low Enrollment**

To help beneficiaries make more informed choices and to be as fully transparent as possible about the performance of all plans, CMS is moving toward including low enrollment contracts in the Star Ratings. Low enrollment contracts, as defined in §422.252, are those where enrollment is such that HEDIS and HOS data collections cannot be used to reliably measure the

performance of the health plan. In the past, we have believed that contracts with less than 1,000 enrollees would meet that definition, but we have reevaluated whether that threshold is an appropriate implementation of the regulatory standard. Contracts with less than 1,000 enrollees first submitted HEDIS data to CMS in the summer of 2013. As a precursor to including low-enrollment contracts in the Star Ratings, CMS included HEDIS scores for low-enrollment contracts as part of the 2014 display measures. For the 2014 Star Ratings, 27 additional contracts would have received an overall rating. Based on the data we received, CMS has determined that there are sufficient data to reliably measure and report on contracts in the Star Ratings with 500 or more enrollees in July of the HEDIS measurement year.

Last year CMS delayed including contracts with enrollment from 500 to 999 enrollees into the Star Ratings on Medicare Plan Finder to gain an additional year of experience with collecting and analyzing these data and to evaluate the reliability of the data. Beginning with the 2016 Star Ratings, contracts with 500 or more enrollees as of July 2014 will be included and used for 2017 QBPs. Contracts with 500 or more enrollees in most cases will have sufficient data to produce both overall and Part C and D ratings. The HEDIS data for contracts with less than 500 enrollees will continue to be posted on the display page as these will continue to be considered low enrollment contracts.

In 2014 CMS provided low enrollment contracts their simulated Star Ratings data and will provide it again in 2015. It is important to note that only the measures where the contract meets the minimum denominator requirements are included in the Star Ratings. Thus, if a contract with 500 to 999 enrollees does not meet the minimum denominator requirements for a measure, the particular measure will not be included in its overall rating calculation. For the HEDIS measures, we plan to exclude from the cut point determinations and the overall rating calculations any contract-specific measure scores that have low reliability. Specifically, any contracts with 500-999 enrollees that have a contract-level reliability of less than 0.7 for a measure will be excluded. The contract-level reliability measures the signal-to-noise ratio which is how much of what is being measured is “signal” (true variation in performance), rather than “noise” (measurement error). Reliability levels of 0.7 or more is considered acceptable reliability.

#### ***H. Data Integrity***

It is essential that the data used for CMS' Star Ratings are accurate and reliable. CMS' policy is to reduce a contract's measure rating to 1 star if it is identified that biased or erroneous data have been submitted. This would include cases where CMS finds mishandling of data, inappropriate processing, or implementation of incorrect practices by the organization/sponsor have resulted in biased or erroneous data. Examples would include, but are not limited to: a contract's failure to adhere to HEDIS, HOS, or CAHPS reporting requirements; a contract's failure to adhere to Plan Finder or PDE data requirements; a contract's errors in processing coverage determinations/exceptions or organization determinations; compliance actions due to errors in operational areas that would directly impact the data reported or processed for specific measures; or a contract's failure to pass Part C and D Reporting Requirements data validation related to organization/sponsor-reported data for specific measures. CMS may perform additional audits or reviews to ensure the

validity of data for specific contracts. Without independent validation of these data, there is risk that CMS will reward contracts with falsely high ratings in these areas.

CMS has taken several steps in the past years to protect the integrity of the data; however, we continue to identify new vulnerabilities where inaccurate or biased data could exist. We are interested in developing more comprehensive quality checks for measures using organization or sponsor-reported data, for example, the Part C and D appeals measures which use data that sponsors report to the IRE. Sponsors have commented in the past that they too are supportive of a comprehensive review of their processes, in lieu of focused or targeted sampling to determine if errors have been made, but at no additional costs to sponsors.

CMS began using validated Part C and D plan reported data for the 2015 Star Ratings with the introduction of the SNP Care Management measure. In order to be evaluated in this measure, contracts must score at least 95% for the SNP Care Management reporting section, and also be found by the data validator to be compliant with data validation standards/sub-standards for the specific data elements used for the measure.

We propose to expand our use of the Part C and D data validation results as a new method of comprehensively reviewing sponsors' operational systems, and verify the validity of the reported data for Star Ratings. Per the Part C and D reporting requirements, contracts submit various data related to their processing of organization determinations, coverage determinations, and appeals, including the timeliness of their processing, data related to the delivery of their MTM programs, etc. Independent data validators assess if these data were accurately reported.

For example, contracts who fail data validation for specific data elements related to organization determination, coverage determination or redetermination timeliness would be found to have biased the data reported to the IRE, and therefore should be reduced in the respective Part C or D appeals star rating measures. Similar applications could be determined for other reporting areas directly relevant to Star rating measures.

We performed an analysis of Part D data reported by sponsors for CY 2013 which were independently validated in April-June 2014. A total of 62 contracts were found to be non-compliant with accurate reporting of timely Coverage Determinations/Exceptions or Redeterminations data (4 contracts were noncompliant in both sections' data). Of these 62, 8 contracts were already reduced by CMS in the corresponding 2015 appeals measures due to CMS program audit findings of CDAG deficiencies. Therefore, our analysis shows that if this method was applied for the 2015 Star Ratings, two times the number of contracts would have been reduced in the Part D appeals Star Rating measures. Since not all sponsors are audited each year, this method may more comprehensively capture evidence of biased data. CMS welcomes feedback to this proposal to expand the data integrity checks to use the Part C and D Data Validation results for associated measures.

## ***I. Dual/LIS Status***

Over the past year, multiple MA organizations and PDP Sponsors have suggested that plans that enroll a disproportionate share of dual-eligible beneficiaries (or low income subsidy (LIS) beneficiaries) may experience difficulty in achieving higher quality care as measured by items included in the MA and Part D Star Ratings Program. In addition, we have reviewed recent information about the impact of socio-economic status on quality ratings, such as the report posted at [www.qualityforum.org/risk\\_adjustment\\_ses.aspx](http://www.qualityforum.org/risk_adjustment_ses.aspx), published by the National Quality Forum (NQF). Because of our interest in using the Star Rating system to foster continuous quality improvement in the MA and Medicare Prescription Drug programs, we are interested to learn whether some plans are truly at a disadvantage. While there is evidence of an association between higher dual-eligible enrollment (and higher LIS beneficiary enrollment) and lower Star Ratings, this association does not prove causality. Indeed, it may be that dual-eligible and LIS beneficiaries are experiencing lower quality care, which would be of paramount concern to CMS. It may also be that dual status is related to other member characteristics such as disability status, chronic conditions, and physical or mental health status.

In a Request for Information (RFI) this fall, CMS sought analyses and research that demonstrate that dual status causes lower MA and Part D quality measure scores, as well as research that demonstrates that high quality performance in MA or Part D plans can be achieved in plans serving dual eligible beneficiaries and how that performance level is obtained. CMS received 67 responses to the RFI. We appreciate the feedback received from multiple organizations. We are currently analyzing the data and information submitted, as well as CMS analyses. If we find that dual/LIS status or other member characteristics are driving performance on particular measures, this information will be shared with the measure developers. More information will be included in the draft Call Letter to be published in February 2015.

## ***J. Measures Posted on the CMS Display Page***

Display measures on [www.cms.gov](http://www.cms.gov) are not part of the Star Ratings. These may include measures that have been transitioned from the Star Ratings, new measures that are tested before inclusion into the Star Ratings, or measures displayed for informational purposes. Similar to the 2015 display page, organizations and sponsors have the opportunity to preview their data on the display measures prior to release on CMS' website. Data on measures moved to the display page will continue to be collected and monitored, and poor scores on display measures are subject to compliance actions by CMS. It is expected that all 2015 display measures will continue to be shown on [www.cms.gov](http://www.cms.gov). CMS will continue to provide advance notice regarding measures considered for implementation as future Star Ratings. Other display measures may be provided as information only.

## ***K. Forecasting to 2017 and Beyond***

The following describes potential changes to existing measures and potential new measures. All of the HEDIS changes and additions are tentative pending a final decision by the NCQA Committee on Performance Measurement and the Board of Directors in June 2015. We also describe potential changes to CAHPS measures to reflect AHRQ's CAHPS 5.0 Health Plan Survey.

## **Potential changes to existing measures:**

### Medication Reconciliation Post Discharge:

The Medication Reconciliation Post-Discharge (MRP) measure assesses the percentage of discharges from acute or non-acute inpatient facilities for members 66 years of age and older for whom medications were reconciled within 30 days of discharge. NCQA is proposing two changes: 1) expand the coverage on this measure from Medicare Special Needs Plans only to all of MA; and 2) expand the age range from adults 65 years and older to adults 18 years and older. Both of these proposed changes for HEDIS 2016 are seen as an important step to measure the quality of care coordination post-discharge for MA beneficiaries as well as ensuring patient safety. If this measure is implemented for HEDIS 2016, CMS will include in the 2017 Display Page and will consider for the 2018 Star Ratings.

### CAHPS measures:

The current MA & PDP CAHPS Survey includes the core CAHPS 4.0 Health Plan Survey. CMS is interested in potentially updating the survey for future years to reflect AHRQ's CAHPS 5.0 Health Plan Survey. We will conduct an experiment in 2015 to understand if/how performance on CAHPS measures differs between 4.0 and 5.0. Based on these results we will consider whether changes or adjustments should be made to the MA & PDP CAHPS Surveys in the future.

### MPF Price Accuracy

CMS is considering updating the MPF Price Accuracy measure in the future. The first proposed change is related to the method in which claims are excluded from the measure. Currently, the measure is limited to 30-day claims filled at pharmacies reported by sponsors as retail only or retail and limited access only in their Medicare Plan Finder (MPF) Pharmacy Cost files. That is, claims filled for near 30 days supplies, or claims filled for 60 and 90 days supplies are excluded. Additionally, claims filled by pharmacies reported to be retail as well as long term care, mail order, or home infusion are excluded. These restrictions result in the exclusion of many PDEs, thus potentially biasing the reliability of the measure.

We propose to include claims with 28-34 days supplies, as we believe it would be appropriate to compare their PDE costs to MPF's fixed display of 1 month pricing. We also propose to include 60 and 90 days supply claims. Beginning with CY2015 MPF submissions, plans must provide brand and generic dispensing fees for 60 and 90 day supply claims in the Pharmacy Cost file. CMS can use these data, along with 60 and 90 days supply Pricing File data, to compare MPF and PDE costs. While the majority of claims are for 30 days supply, we found that claims with 90 days supply account for almost one-fifth of available PDE data, thus allowing for a more comprehensive evaluation of PDE claims.

Additionally, we propose to use the PDE-reported Pharmacy Service Type code in conjunction with the MPF Pharmacy Cost data to identify retail claims. Prior to the availability of this PDE field, there was no way to determine whether a given claim was priced under the retail setting of the dispensing pharmacy when a pharmacy had multiple types. There may be incentives for

sponsors to misreport pharmacy types in the MPF Pharmacy Cost files to reduce the number of claims eligible for inclusion in the Price Accuracy Score. CMS began requiring pharmacies to populate the Pharmacy Service Type field on all PDEs at the end of February 2013. As of June 2014, the Pharmacy Service Type field was populated for 99.9 percent of CY2014 PDEs submitted. We recommend expanding the retail claims identification process to include all PDEs that are from at least retail pharmacies according to the Pharmacy Cost data and have a Pharmacy Service Type of either Community/Retail or Managed Care Organization (MCO). This methodology change would increase the number of PDEs eligible for inclusion in the Price Accuracy Scores while continuing to identify only retail claims.

These proposed changes can also be applied to mail order claims. Including mail order claims with 28-34, 60, and 90 days supplies would add another dimension to the Price Accuracy Scores and further increase the number of PDEs eligible for inclusion. CMS can use the Pharmacy Cost data to identify PDEs from pharmacies that are at least mail order, and then restrict to PDEs that are Mail Order according to the Pharmacy Service Type field. Then, using Pricing File data for 30, 60, and 90 days supply mail order claims, as well as brand and generic dispensing fees from the Pharmacy Cost data, we can compare MPF and PDE costs for mail order claims.

We are also considering changes to the methodology by which price accuracy is calculated. Because the current methodology measures the magnitude of a contract's overpricing relative to its overall PDE costs, the Price Accuracy Scores do not reflect the frequency of accurate price reporting, and can be significantly impacted by high cost PDEs. As a result, contracts with divergent accurate price reporting and/or consistency can receive the same Price Accuracy Score. CMS is interested in modifying the methodology to also factor in how often PDE costs exceeded MPF costs. The frequency of inaccuracy by a contract would be the percent of claims where PDE cost is greater than MPF cost. The numerator is the number of claims where PDE cost is greater than MPF cost, and the denominator is the total number of claims. This ratio is then subtracted from 1 and multiplied by 100 to calculate the Claim Percentage Score, with 100 as the best possible score and 0 as the worst possible score. The contract's accuracy score would be a composite of the Price Accuracy Score and the Claim Percentage Score. By capturing the frequency of inaccuracy as well as the magnitude, the measure would better depict the reliability of a contract's MPF advertised prices. Other options we explored included measuring the magnitude of inaccuracy as a percentage cost difference, instead of the current measure's use of absolute cost difference. Testing however found this method may overstate small differences between PDE and MPF costs for low-cost claims. For example, when using percentage cost differences, a claim with a \$2.00 PDE cost and a \$1.00 MPF cost would be considered equally overpriced as a claim with a \$200.00 PDE cost and a \$100.00 MPF cost.

We believe the proposed changes will greatly improve the Price Accuracy Scores, making them a more comprehensive assessment of contracts' price reporting for Part D beneficiaries.

## **Potential new measures:**

### Care Coordination Measures:

Effective care coordination contributes to improved health outcomes. CMS believes that 5-star plans perform well on our Star Ratings measures because they understand how to effectively coordinate care for their enrollees. Our assumption about plans, however, is based largely on anecdote and discussions with high-performing plans, as we currently lack the tools to accurately capture and measure how well plans are coordinating care.

To date, our ability to measure plans' care coordination efforts has largely been limited to data we collect from CAHPS surveys, which reflect enrollees' experience with the care they receive. CMS is working to expand efforts in this area to measure the plans' coordination approaches. These efforts will focus on developing measures related to the patient assessment of their plans' care coordination, encounter data-based measures, and medical records-based measures. CMS is particularly interested in comments on measures that could be developed using MA encounter data. For example, measures that identify post-discharge utilization by plan enrollees in order to identify plans in which an unusually high number (proportion) of enrollees do not obtain expected follow-up care (follow-up physician visit within first week) or, if appropriate, for whom there are no changes to prescribed medications following discharge. As measures are developed and tested, they will be added to the Display Page and Star Ratings.

CMS will also monitor any additional measures developed by NCQA or PQA for potential incorporation into the Star Ratings.

### Asthma Measure Suite:

NCQA will be testing three asthma measures in the fall of 2014 to evaluate the effects of expanding the measure to include older adults. The age range for these measures is currently members 5 – 64 years of age. The three measures under consideration for inclusion of older adults include:

- Use of Appropriate Medications for People with Asthma: The percentage of members during the measurement year who were identified as having persistent asthma and who were appropriately prescribed medication during the measurement year.
- Medication Management for People with Asthma: The percentage of who were identified as having persistent asthma and were dispensed appropriate medications that they remained on during the treatment period (i.e. first prescription date through end of measurement year).
- Asthma Medication Ratio: The percentage of members who were identified as having persistent asthma and had a ratio of controller medications to total asthma medications of 0.50 or greater during the measurement year.

Testing results will be reviewed with NCQA's measurement advisory panels, including the Geriatric Measurement Advisory Panel. These panels will help NCQA determine whether

expanding the age range of these measures to include the 65+ population is appropriate. The proposed changes, if approved, would be published in HEDIS 2016.

#### Depression:

NCQA is developing a new set of HEDIS measures that would assess depression care along the continuum of care. These measures are intended for all individuals age 12 and older but may be particularly relevant to the population age 65 and older. The measures currently in testing include:

- Depression Screening and Follow-up: The percentage of individuals who were screened for depression using a standardized tool and received appropriate follow-up in the screen was positive
- Utilization of the PHQ9 for Monitoring of Depressive Symptoms: The percentage of individuals with a diagnosis of major depression or dysthymia who were monitored using the Patient Health Questionnaire (PHQ-9).
- Depression Remission, Response or Treatment Adjustment at 6 Months: The percentage of individuals with a diagnosis of major depression or dysthymia and symptomatic depression at baseline who achieved either remission of depression symptoms, response (i.e. reduction) in symptoms or an adjustment in treatment at six months.

#### Hospitalizations for Potentially Preventable Complications:

NCQA is finalizing testing of a risk-adjusted measure of hospitalization for ambulatory care sensitive conditions based on the NQF endorsed Prevention Quality Indicators (PQI), developed by AHRQ. This measure will assess the rate of hospitalization for complications of chronic and acute ambulatory care sensitive conditions. The intent of the measure is to assess the quality of ambulatory care to prevent the complications of chronic and acute conditions that result in hospitalization. The new measure, if approved, would be published in HEDIS 2016.

#### Statin Therapy:

NCQA is currently developing two statin therapy measures aligned with the 2013 ACC/AHA blood cholesterol guidelines. The measures are focused on two of the major statin benefit groups described in the guidelines: patients with clinical atherosclerotic cardiovascular disease and patients with diabetes. Measure development and field-testing is expected to continue through the fall and winter. The new measures if approved would be published in HEDIS 2016.

The PQA is developing a new measure to support ACC/AHA guidelines which recommend moderate- to high- intensity statin therapy for primary prevention for patients aged 40-75 years of age with diabetes. The measure calculates the percentage of patients ages 40 – 75 years who received a medication for diabetes that receive a statin medication during the measurement period. This measure may be considered for endorsement by the PQA membership as early as November 2014, so CMS is closely following any such decisions. Once endorsed, CMS will continue to test this measure, explore developing new reports to Part D

sponsors via the Patient Safety Analysis website, and evaluate adding this measure as a future Part D Star Rating. For example, if this measure is endorsed by the PQA in 2014, this measure could be considered as a new 2017 Display Measure (using 2015 data) and a 2018 Star Rating (using 2016 data). Patient safety reports to sponsors may be released as early as 2015.

#### High Risk Medication (HRM):

The American Geriatric Society (AGS) is currently considering revisions to the Beer's criteria which may precipitate future changes to the PQA measure specifications and medication list. CMS is closely following these activities. If changes are published by the AGS and measure updates endorsed by the PQA prior to the 2016 bid deadline in June 2015, CMS may consider adoption for the 2018 Star Ratings (using 2016 data).

#### Opioid Overutilization:

PQA is currently developing three measures that examine multi-provider, high dosage opioid use among individuals 18 years and older without cancer. Patients enrolled in hospice are also excluded. The measures currently in development include:

- Measure 1 (Opioid High Dosage): The proportion (XX out of 1,000) of individuals without cancer or enrolled in hospice receiving a daily dosage of opioids greater than 120mg morphine equivalent dose (MED) for 90 consecutive days or longer.
- Measure 2 (Multiple Prescribers and Multiple Pharmacies): The proportion (XX out of 1,000) of individuals without cancer or enrolled in hospice receiving prescriptions for opioids from four (4) or more prescribers AND four (4) or more pharmacies.
- Measure 3 (Multi-Provider, High Dosage): The proportion (XX out of 1,000) of individuals without cancer or enrolled in hospice receiving prescriptions for opioids greater than 120mg morphine equivalent dose (MED) for 90 consecutive days or longer, AND who received opioid prescriptions from four (4) or more prescribers AND four (4) or more pharmacies.

If these measures are endorsed by the PQA prior to the 2016 bid deadline in June 2015, CMS may consider adoption as display measures beginning in CY 2015 and for the 2018 Star Ratings (using 2016 data) or for use in the Overutilization Monitoring System (OMS).

#### **L. Measurement Concepts**

CMS is committed to continuing to improve the Part C and D Star Ratings by identifying new measures and methodological enhancements. Feedback or recommendations can help CMS' continuing analyses, as well as our collaboration with measurement development entities such as NCQA and PQA.

- We welcome comments and input on alternative levels of evaluation (e.g., PBP or parent organization). We are specifically interested in understanding how physician and provider networks may differ across PBPs resulting in differences in quality across PBPs.
- We also welcome comments on additional measures (e.g., care coordination, care transitions, patient-reported outcomes/intermediate outcomes collected through

enrollee surveys, condition-specific measures, SNP-specific measures, measures for the disabled, and outcomes based measures for MTM).

- We have also heard concerns from sponsors with low enrollment that certain measures (such as complaints and appeals) and associated star assignments may be sensitive to small measure denominator size. For the 2016 Star Ratings, we proposed to expand the measurement period for the Complaints about the Health Plan/Drug Plan measures and the Appeals Upheld (Part D) measure to 12 months to increase the number of enrollees included in these measures. This should help minimize any potential concerns, but we welcome feedback from sponsors on this matter, including analysis reflecting any sensitivity issues and potential solutions.
- In addition, we are interested in feedback to whether organization-specific cut points are relevant for some Part D measures, when organization type should not result in performance differences. For example, the MPF price accuracy measure evaluates differences in a Part D Sponsors' submitted MPF price files and PDE files, and CMS questions if MA-PDs and PDPs are comparable for this process measure and could be measured together.

## Appendix A

### Measure – Beneficiary Access and Performance Problems (Revised Methodology)

**Labels for Stars:** Problems Medicare Found in Members' Access to Services and in the Plan's Performance (more stars are better because it means fewer serious problems)

**Label for Data:** Problems Medicare Found in Members' Access to Services and in the Plan's Performance (on a scale from 0 to 100, **higher numbers** are better because it means fewer serious problems)

**Description:** To check on whether members are having problems getting access to services and to be sure that plans are following all of Medicare's rules, Medicare conducts several different types of reviews. Medicare gives the plan a **lower** score (from 0 to 100) when it finds problems. The score combines **how severe** the problems were, **how many** there were, and **how much** they affect plan members directly. A higher score is better, as it means Medicare found fewer problems.

**Metric:** This measure is based on CMS' sanctions, civil monetary penalties (CMP) as well as Compliance Activity Module (CAM) data (this includes: notices of non-compliance, warning letters {with or without business plan}, and ad-hoc corrective action plans (CAP) and the CAP severity).

- Contracts' scores are based on a scale of 0 -100 points.
- The starting score for each contract works as follows:
  - Contracts with an effective date of 1/1/2014 or later are marked as "Plan too new to be measured".
  - All contracts with an effective date prior to 1/1/2014 begin with a score of 100.
- Contracts placed under sanction anytime during the data time frame are reduced to a score of 0. This is separate from the deduction applied at the overall score level for contracts with more recent sanctions.
- The following deductions are taken from contracts whose score is above:  
For each CMP, Contracts that received a CMP with beneficiary impact related to access: 40 points.  
Contracts that have a CAM score (CAM score calculation is discussed below) are reduced as follows:
  - 0– 2 CAM Score – 0 points
  - 3 – 9 CAM Score – 20 points
  - 1 – 19 CAM Score – 40 points
  - 2 – 29 CAM Score – 60 points
  - ≥ 3 CAM Score – 80 points

Calculation of the CAM Score combines the notices of noncompliance, warning letters (with or without business plan) and ad-hoc CAPs and their severity. The formula used is as follows:

$$\text{CAM Score} = (\text{NC} * 1) + (\text{woBP} * 3) + (\text{wBP} * 4) + (\text{NAHC} * (6 * \text{CAP Severity}))$$

Where: NC = Number of Notices of Non Compliance

woBP = Number of Warning Letters without Business Plan

wBP = Number of Warning Letters with Business Plan

NAHC = Number of Ad-Hoc CAPs

CAP Severity = Sum of the severity of each individual ad-hoc CAP given to a contract during the measurement period. Each CAP is rated as one of the following:

- 3 – ad-hoc CAP with beneficiary access impact
- 2 – ad-hoc CAP with beneficiary non-access impact
- 1 – ad-hoc CAP no beneficiary impact

**Data Source:** CMS Administrative Data

**Data Source Description:** Findings of CMS compliance actions that occurred during the 12 month past performance review period between January 1, 2014 and December 31, 2014. For compliance actions, the date the action was issued is used when pulling the data from HPMS.

**CMS Framework Area:** Population/Community Health

**NQF#:** None

**Data Time Frame:** 01/01/2014 – 12/31/2014

**General Trend:** Higher is better

**Statistical Method:** Relative Distribution and Clustering

**Improvement Measure:** Not Included

**Weighting Category:** 1.5

**Data Display:** Rate with no decimal point

**Reporting Requirements:**

1876 Cost	Demo	Local, E-local, RPPO, CCP w/o SNP	Local, E-local, RPPO, CCP w/SNP	MSA	E-PDP & PDP	E-PFFS, PFFS
Yes	Yes	Yes	Yes	Yes	No	Yes

**4- Star threshold:** Not predetermined

Appendix B

**Improvement measures (Part C & D):**

Part	Measure	Improvement Measure
C	Breast Cancer Screening	No
C	Colorectal Cancer Screening	Yes
C	Annual Flu Vaccine	Yes
C	Improving or Maintaining Physical Health	No
C	Improving or Maintaining Mental Health	No
C	Monitoring Physical Activity	Yes
C	Adult BMI Assessment	Yes
C	Special Needs Plan (SNP) Care Management	Yes
C	Care for Older Adults – Medication Review	Yes
C	Care for Older Adults – Functional Status Assessment	Yes
C	Care for Older Adults – Pain Assessment	Yes
C	Osteoporosis Management in Women who had a Fracture	Yes
C	Diabetes Care – Eye Exam	Yes
C	Diabetes Care – Kidney Disease Monitoring	Yes
C	Diabetes Care – Blood Sugar Controlled	Yes
C	Controlling Blood Pressure	Yes
C	Rheumatoid Arthritis Management	Yes
C	Reducing the Risk of Falling	Yes
C	Plan All-Cause Readmissions	Yes
C	Getting Needed Care	Yes
C	Getting Appointments and Care Quickly	Yes
C	Customer Service	Yes
C	Rating of Health Care Quality	Yes
C	Rating of Health Plan	Yes
C	Care Coordination	Yes
C	Complaints about the Health Plan	Yes
C	Members Choosing to Leave the Plan	Yes
C	Beneficiary Access and Performance Problems	No
C	Health Plan Quality Improvement	No
C	Plan Makes Timely Decisions about Appeals	No
C	Reviewing Appeals Decisions	Yes
C	Call Center – Foreign Language Interpreter and TTY Availability	No
D	Call Center – Foreign Language Interpreter and TTY Availability	No
D	Appeals Auto-Forward	Yes
D	Appeals Upheld	No
D	Complaints about the Drug Plan	No
D	Members Choosing to Leave the Plan	Yes
D	Beneficiary Access and Performance Problems	No
D	Drug Plan Quality Improvement	No
D	Rating of Drug Plan	Yes
D	Getting Needed Prescription Drugs	Yes
D	MPF Price Accuracy	No
D	High Risk Medication	Yes
D	Diabetes Treatment	Yes
D	Medication Adherence for Diabetes Medications	No
D	Medication Adherence for Hypertension (RAS antagonists)	Yes
D	Medication Adherence for Cholesterol (Statins)	Yes
D	Medication Therapy Management Program Completion Rate for Comprehensive Medication Reviews	No