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

CENTER FOR MEDICARE

DATE: November 12, 2015

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 2017 and Beyond (Due by December 10, 2015 at 5pm ET)

This document proposes methodology changes for the 2017 Star Ratings and display measures for Medicare Advantage (MA) and Prescription Drug Plans (PDPs). It also provides advanced notice of potential changes for the Star Ratings and display measures for 2018 and beyond. Based on this memo, MA Organizations, PDP sponsors, advocates, and other stakeholders have the opportunity to provide comments prior to the draft Call Letter, which is issued as part of the Advance Notice. The statutory deadlines in section 1853 result in a short comment period between the Advance and Final Rate Notices, which contain the Call Letter. We are issuing this Request for Comments (RFC) to provide an opportunity for stakeholders to provide input 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 true outcomes;
- 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-of-care 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 (MPF) 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 and focuses it on beneficiary outcomes, beneficiary experience, 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. Below is an example for the next cycle of Star Ratings, beginning with this RFC:

- November 2015-February 2016: CMS provides guidance on methodology changes anticipated for the 2017 Star Ratings and display measures and advanced notice of potential changes for the Star Ratings and display measures for 2018 and beyond. CMS does this first through the RFC and then in the draft 2017 Call Letter. Sponsors and other stakeholders actively participate by submitting comments regarding the potential changes to the Star Ratings methodology or submitting additional comments to propose other changes not mentioned in the RFC. Comments to the RFC inform proposals in the draft Call Letter, and comments to the draft Call Letter inform the final Call Letter.
- February-April 2016: After consideration of all comments, CMS announces the methodology for the 2017 Star Ratings in the final 2017 Call Letter in the Rate Announcement issued.
- Ongoing: Sponsors review the underlying data used for the individual Star Ratings measures as they become available throughout the year, and notify CMS of any errors or questions in a timely manner.
- Summer 2016: CMS holds a Part C and D User call with sponsors to provide updates to the upcoming Star Ratings release, and conducts two plan preview periods via the Health Plan Management System (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.
- November 2016: 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 non-methodological challenges, such as a calculation error (miscalculation) or a data inaccuracy (incorrect data).

Comments to this RFC should be submitted via the following link:
https://cmsgov.wufoo.com/forms/enhancements-to-the-star-ratings-for-2017/

The online form allows comments on up to 8 sections. If your organization's comments exceed this maximum, you may submit the form more than once. Do not resubmit these comments to CMS via email. If you wish to submit additional supporting documents, you may send them via email to: PartCandDStarRatings@cms.hhs.gov.

NOTE: If you encounter the following error message while completing the survey, you must clear cookies and then resubmit the form: "There was a problem with your submission. Unable to create a new entry."

Comments submitted by Thursday, December 10 at 5pm ET will be considered as we finalize proposed changes for the 2017 Star Ratings for the draft 2017 Call Letter. CMS will post on its website all comments received to the RFC. Stakeholders will have another opportunity to comment on the 2017 Star Ratings methodology and proposed changes through the Advance Notice/draft Call Letter process.

Questions related to this RFC may be sent to: PartCandDStarRatings@cms.hhs.gov.

Thank you for your participation.

Enhancements to the 2017 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 2017 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 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 2017 Star Ratings and beyond. CMS is not considering adding any new measures for 2017 Star Ratings. Unless noted below, we anticipate the methodology remaining the same as the 2016 Star Ratings.

For reference, the list of measures and methodology included in the 2016 Star Ratings is described in the Technical Notes available on the CMS webpage: http://go.cms.gov/partcanddstarratings.

The cut points to determine star assignments 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 2017 using the most current data available.

As announced in previous years, we will 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.

A. Changes to Measures for 2017

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 announced during the measurement period 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 announced during the measurement period does not significantly impact the numerator or denominator of the measure, the measure will continue to be included in the Star Ratings (e.g., 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. **Improvement measures (Part C & D).** The methodology for the improvement measures remains the same as in prior years. We have updated the measures included in the improvement measure to account for measures with at least two years of data. Please refer to the Appendix for updates to the measures to be used to calculate the 2017 improvement measures. We are also considering that if a contract's CAHPS measure score moved to very low reliability with the exclusion of the enrollees with less than 6 months of continuous

enrollment for the 2015 survey administration, then the 2014 CAHPS measure score (used in 2015 Star Ratings) would be used instead as the baseline for the 2017 improvement calculation for that measure.

2. **Reviewing Appeals Decisions/Appeals Upheld measures (Part C & D).** Currently, these measures include cases that are reopened and decided by April 1 of the following contract year. In some instances, appeals filed in the 4th quarter of the year and then subsequently reopened may not be determined by the Independent Review Entity (IRE) by April 1. We propose for the 2017 Star Ratings to modify these measure specifications so that if a Reopening occurs and is decided prior to May 1, 2016, the Reopened decision would be used. Reopenings decided on or after May 1, 2016 would not be reflected in these data, and the original decision result would be used.

3. **Contract Enrollment Data (Part C & D)**. Contract enrollment numbers are pulled from HPMS for the Part C and D "Complaints about the Health/Drug Plan" and the Part D "Appeals Auto-Forward" measures. Additionally, plan-level enrollment is pulled for the three Part C "Care for Older Adults" measures. For these measures, twelve months of enrollment files are pulled from HPMS, and the average enrollment from those months is used in the measure calculations. We propose going forward to adjust the twelve months from January to December to February through January of the relevant measurement period. HPMS enrollment numbers are determined between the end of the first week to the beginning of the second week of the prior month. For example, January enrollment numbers reflect enrollment as determined as of the first week of December. Thus, we are modifying the enrollment numbers to reflect an average of the HPMS enrollment from February through January of the measurement year.

4. **Transition from ICD-9 to ICD-10 (Part C & D).** The measure stewards, such as the Pharmacy Quality Alliance (PQA) and the National Committee for Quality Assurance (NCQA), are in the process of reviewing their measure specifications with diagnosis-related requirements to transition from ICD-9 to ICD-10.

   Once the PQA updates their measure specifications, we will provide more information. We will test and adopt the changes implemented by PQA as appropriate for the Part D Star Ratings and display measures.

   NCQA has incorporated the ICD-10 codes in the 2016 Healthcare Effectiveness Data and Information Set (HEDIS). During the transition period both ICD-9 and ICD-10 codes will be used due to the look-back periods for some measures.

5. **Appeals Upheld measure (Part D).** For the 2016 Star Rating Upheld measure, we excluded appeal cases for beneficiaries enrolled in hospice at any point during 2014. As noted in the 2016 Call Letter, this exclusion was only necessary for the 2016 measure as it is based on 2014 data that may have been affected by policy changes in 2014. This exclusion will not be continued for the 2017 Star Rating Appeals Upheld measure.

6. **Medication Therapy Management (MTM) Program Completion Rate for Comprehensive Medication Reviews (CMR) measure (Part D).** We will add a detailed file during each HPMS plan preview period to list each contract's underlying denominator, numerator, and Data Validation score since exclusions are applied to the plan-reported MTM data.

   The CMR rate measure is an initial measure of the delivery of MTM services, and we continue to look forward to the development and endorsement of outcomes-based MTM measures as potential companion measures to the current MTM Star Rating. Lastly, we will be implementing additional data integrity checks (discussed later in this memo) to safeguard against inappropriate attempts to bias the data used for this measure.

B. Removal of Measures from Star Ratings

1. **Improving Bladder Control (Part C).** This measure, collected through the Health Outcomes 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 made 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 completed.

   These changes required revising the underlying survey questions in HOS. The revised questions were first collected in 2015. As a result of these changes, this measure will not be reported in the 2017 Star Ratings. The revised measure will be reported on the 2017 display page since the survey was first fielded with the new questions in 2015. The 2016 display measure uses data from the old questions.

2. ***High Risk Medication (Part D).*** The High Risk Medication (HRM) measure calculates the percent of Medicare Part D beneficiaries 65 and older who received two or more prescription fills for the same HRM drug with a high risk of serious side effects in the elderly. The measure is endorsed by the PQA and National Quality Forum (NQF), and the HRM rate is calculated using the PQA specifications and medication list based on American Geriatrics Society (AGS) recommendations. The AGS recently released the 2015 update of the Beers Criteria.

   The HRM measure will be removed from the Star Ratings and moved to the display measures for 2017. This proposal is based on a number of factors. While the AGS states that the criteria may be used as both an educational tool and quality measure, the AGS further states that the intent is not to apply the criteria in a punitive manner. Specifically, the addition of a drug to the HRM list is not a contraindication to use, rather an encouragement to avoid use in

the senior population without consideration of risks and benefits based on individual patient considerations. This is a very difficult decisional balance to evaluate in a drug plan that does not have access to full clinical information. As the measure can be calculated only by using prescription drug event (PDE) data, medications cannot be included on the HRM List that have risks conditional on clinical factors that cannot be measured using PDE data alone. As a result, some "Avoid" medications are included in the measure, while others are not. This may create unintended consequences including the inappropriate encouragement of certain non-HRM medications, which may not be the best choice for an individual beneficiary's clinical circumstance.

Lastly, because it is under direct provider control and should not be affected by non-clinical beneficiary characteristics, the HRM measure was not included in CMS' overall analysis to assess the impact of socio-economic status (SES) on the Star Ratings (discussed later in this memo). However, our initial analysis found that after controlling for contract effects and dual eligible or low income subsidy status, there is a significant association between dual eligible/low income status and HRM use. This association remains after further controlling for age, sex, and race/ethnicity. We recommend that the measure developers further review this measure to better understand the associations.

Avoiding potentially inappropriate medications in older adults remains important for quality of care for Medicare beneficiaries. Therefore, the HRM measure will move to the 2017 display page and may be considered for the Star Ratings again in the future. We will continue to provide HRM measure reports to Part D sponsors on a monthly basis through the Patient Safety Analysis website, and we will continue to identify outliers.

If measure updates are endorsed by the PQA with sufficient lead time ahead of the 2017 formulary and bid deadlines in May and June 2016, CMS may consider adoption for the 2019 display page (using 2017 data). We will provide additional information on updates if available in the draft 2017 Call Letter.

C. 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 determined 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 found through program audits or other reviews; 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 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 also must safeguard against the Star Ratings program creating perverse incentives for sponsors. CMS program audits will soon include review of Part D sponsors' MTM programs. We intend to review and apply any relevant MTM program audit findings that could demonstrate sponsors' MTM data were biased, outside of the Data Validation results. CMS is concerned about sponsor activities that may not be detected by routine Data Validation standards, such as attempts to restrict eligibility from their approved MTM programs, encouraging beneficiary opt-out of MTM programs within the first 60-days, or CMRs that do not meet CMS' definition per guidance. CMS may perform additional audits or reviews to ensure the validity of data for specific contracts. Without rigorous validation of Star Ratings data, there is risk that CMS will reward contracts with falsely high ratings.

D. Impact of Socio-economic and Disability Status on Star Ratings

A key goal of the MA and Part D programs is to achieve greater value and quality for all beneficiaries; therefore, an important corollary is that we do not distort quality signals in our measures, or mask true differences in quality of care. CMS continuously reviews the Star Ratings methodology to improve the process, incentivize plans, and provide information that is a true reflection of the performance and experience of the enrollees. The policies implemented must result in high quality of care and health outcomes for all of our beneficiaries, while acknowledging the unique challenges of serving traditionally underserved subsets of the population.

A number of MA organizations and PDP sponsors believe that enrollment of a high percentage of dual eligible (DE) enrollees and/or enrollees who receive a low income subsidy (LIS) limits their plan's ability to achieve high MA or Part D Star Ratings. CMS has responded to the concern of our stakeholders and has comprehensively gathered information to determine if the Star Ratings are sensitive to the socio-economic and disability status of a contract's enrollees. If adjustments are made to address this issue, they must be data driven. For example, if a disparity is due to challenges in serving disabled beneficiaries, rather than in serving those with lower SES, then the adjustment should clearly focus on disability status of beneficiaries. Similarly, unless our methods are transparent and open to input from a breadth of sources, contracts will not be able to translate as easily our findings into actionable quality improvement steps.

With support from our contractors, we have undertaken research to provide the scientific evidence as to whether MA or Part D sponsors that enroll a disproportionate number of vulnerable beneficiaries are systematically disadvantaged by the current Star Ratings. Last year, we issued a Request for Information to gather information directly from organizations to supplement the data that CMS collects, as we believe that plans and sponsors are uniquely positioned to provide both qualitative and quantitative information that is not available from other sources. In February and September 2015, we released details on our research and

findings to-date.[1] We have also reviewed reports about the impact of SES on quality ratings, such as the report published by the National Quality Forum (NQF) posted at www.qualityforum.org/risk_adjustment_ses.aspx and both the Medicare Payment Advisory Commission's (MedPAC) *Report to the Congress: Medicare Payment Policy* posted at http://www.medpac.gov/documents/reports/mar2015_entirereport_revised.pdf?sfvrsn=0 and their recent release on September 10th entitled *Factors Affecting Variation in Medicare Advantage Plan Star Ratings* posted at http://www.medpac.gov/documents/september-2015-meeting-presentation-factors-affecting-variation-in-medicare-advantage-plan-star-ratings.pdf?sfvrsn=0. The IMPACT Act (P.L. 113-185) instructs ASPE (Office of the Assistant Secretary for Planning and Evaluation) to conduct a study before October 2016 that examines the effect of individuals' SES on quality measures, resource use and other measures for individuals under the Medicare program. Because ASPE's research agenda aligns closely with our goals, we have and will continue to work collaboratively with ASPE and other governmental agencies to broaden and expand the focus of the issue. We note that, as instructed by Congress in the IMPACT Act, ASPE is conducting additional research in this area and may make recommendations for additional changes in the future. We look forward to their continued input. Further, CMS has engaged measure developers, NCQA and the PQA, to examine measure specifications used in the Star Ratings program to determine if measure re-specification is warranted.

CMS' work is not complete, and we will continue to work diligently to address this issue and others that may lie in the future with the goal that all MA and Part D beneficiaries receive the highest quality care possible. The Star Rating system was designed to foster continuous quality improvement in the MA and Part D programs. As such, the Star Ratings program - its measures and methodology - are reviewed on an ongoing basis. We are committed to providing beneficiaries information on Medicare Plan Finder that is a true reflection of the care and experience of the plans' members and to incentivize plans based on this same information.

As stated in the 2016 Final Call Letter, CMS believed additional research into the nature of the differential performance on a subset of measures was necessary before any interim or permanent changes in the Star Ratings measurements could be developed and implemented.

The additional research conducted after the publication of the 2016 Final Call Letter allowed for further examination of LIS/DE differences ("effects") and their magnitude. Due to the considerable overlap between LIS/DE beneficiaries and disabled beneficiaries, the research was expanded to consider the possible role of disability status. The research considered the association between the performance on Star Ratings measures and enrollment of LIS/DE/disabled beneficiaries, and the variability of differences in performance on each measure by contract to

[1] The February release can be found at https://www.cms.gov/medicare/prescription-drug-coverage/prescriptiondrugcovgenin/performancedata.html
The September release can be found at https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/Research-on-the-Impact-of-Socioeconomic-Status-on-Star-Ratingsv1-09082015.pdf

gain a better understanding of LIS/DE differences revealed in the preliminary research.[2] The methodology employed allowed for the delineation of within- and between-contract differences associated with LIS/DE and/or disability. Within-contract differences are differences that may exist between subgroups of enrollees in the same contract (e.g., if LIS/DE enrollees within a contract have a different mean or average performance on a measure than non-LIS/DE enrollees in the same contract). These differences may be favorable or unfavorable for LIS/DE/disabled beneficiaries and can be assessed separately from the overall level of performance for a contract. Between-contract differences in performance associated with LIS/DE/disability status ("between-contract disparities") are the possible additional differences in performance between contracts associated with the contract's proportion of LIS/DE/disabled enrollees that remain after considering within-contract disparities by LIS/DE/disability status. If LIS/DE/disabled beneficiaries are more or less likely than other beneficiaries to be enrolled in lower-quality contracts, then between-contract disparities may represent true differences between contracts in quality. Because of this possibility, between-contract disparities may not be appropriate for adjustment due to the risk of masking true differences in quality. Adjusting for within-contract disparities is an approach aligned with the consensus reflected in the NQF report on sociodemographic adjustment, which states that, "*…only the within-unit effects are adjusted for in a risk adjustment procedure because these are the ones that are related specifically to patient characteristics rather than differences across units*" (National Quality Forum, 2014). Our research focused on measuring within-contract differences in performance for LIS/DE/disabled compared to non-LIS/DE/disabled beneficiaries.

Our additional research findings are consistent with the preliminary results shared in the 2016 Final Call Letter. The research to date has provided scientific evidence that there exists a within-contract LIS/DE/disability effect for a subset of the Star Ratings measures. The size of the effect differs across measures and is not exclusively negative.

CMS is firmly committed to building the foundation for a long-term solution that appropriately addresses the issue at hand and aligns with our policy goals. Any policy response must delineate the two distinct aspects of the LIS/DE and/or disability issue - quality and payment. The Star Rating Program focuses on accurately measuring the quality of care provided so any response must focus on enhancing the ability to measure actual quality differences among contracts. To address the LIS/DE/disability issue we must accurately address any sensitivity of the ratings to the SES of the beneficiaries enrolled in a contract at the basic building block of the rating system, the measure. CMS has encouraged the measure stewards to examine our findings and undertake an independent evaluation of the measures' specifications to determine if measure re-specification is warranted. Concurrently, the payment response must focus on resource utilization and the predictive performance of the risk-adjustment models for

[2] The research focused on a total of 16 clinical quality measures. A measure was *excluded* from analysis if the measure was already case-mix adjusted for SES (i.e., CAHPS and HOS measures), the focus of the measurement was not a beneficiary-level issue but rather a plan-level issue (e.g., appeals, call center, Part D price accuracy), the measure was scheduled to be retired or revised, or the measure was applicable to only Special Needs Plans (SNPs) (i.e., SNP Care Management, Care for Older Adults measures).

the unique cost patterns of beneficiaries in the community. CMS is also considering changes in the risk adjustment models for payment and issued a separate Request for Comments on October 28, 2015 to obtain feedback on potential revisions. We feel that these two approaches are complementary; holding contracts to a same quality standard is most appropriate when contracts are adequately resourced to provide the support their beneficiaries need to achieve good health outcomes.

While the measure stewards are undertaking a comprehensive review of their measures used in the Star Ratings program, CMS is exploring two options for interim analytical adjustments to address the LIS/DE/disability effect: a Categorical Adjustment Index or Indirect Standardization. We believe each of the proposed methods, discussed in more detail below, align with the goals of: making adjustments that reflect the actual magnitude of the differences observed in the data; providing valid quality ratings to facilitate consumer choice; and providing incentives for MA and Part D quality improvement. In addition, we recognize the need for the options to be both transparent and feasible for the plans, as well as to maintain the integrity of the Star Ratings and the core of its methodology.

Another issue we are examining is the manner to address the unique aspects of implementation of Medicare in Puerto Rico. Under statute, many of Department of Health and Human Services' (HHS) programs, including Medicare and Medicaid, are implemented differently in Puerto Rico. In addition, Puerto Rico has a unique health care market with many low-income individuals in both Medicare and Medicaid and a complex legal history that affects the health care system in many ways. We are cognizant of the particular challenges in Puerto Rico and propose an additional analytical adjustment for contracts with a service area only in Puerto Rico to address the fact that the Part D low income subsidy (LIS) is not available there.

The details of the possible interim analytical adjustments follow. While the interim policy responses to address the LIS/DE/disability effect are distinct analytical adjustments that are applied on the measure scores, they offer flexibility in their application. CMS is requesting feedback on each of the possible adjustments and the possible permutations of each option, or potential hybrid approaches. We welcome comments on the specific measures that should be adjusted[3].

The Categorical Adjustment Index

The Categorical Adjustment Index is a factor that would be added or subtracted to a contract's Overall and/or Summary Star Rating to adjust for the average within-contract disparity. Contracts would be categorized based on their percentages of LIS/DE and/or disabled

[3] Our research focused on the following 16 measures: adult BMI assessment, rheumatoid arthritis management, breast cancer screening, controlling blood pressure, diabetes care – blood sugar controlled, diabetes care – eye exam, diabetes care – kidney disease monitoring, colorectal cancer screening, osteoporosis management in women who had a fracture, plan all-cause readmissions, annual flu vaccine, monitoring physical activity, reducing the risk of falling, medication adherence for diabetes medications, medication adherence for hypertension, and medication adherence for cholesterol.

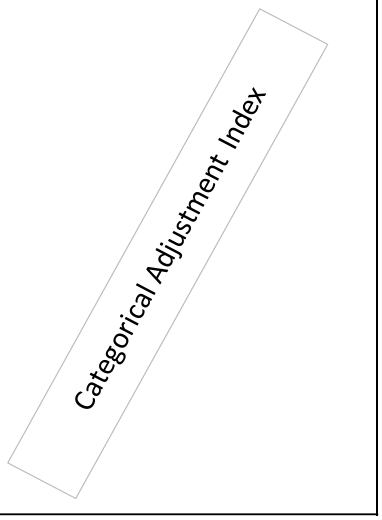
beneficiaries, and the Categorical Adjustment Index value would be the same for all contracts within each category. The Categorical Adjustment Index value is the star adjustment for contracts by category.

The Categorical Adjustment Index values would be computed by comparing the mean Overall and/or Summary contract Star Rating derived from measure scores that are adjusted for LIS/DE and/or disability status to the mean Star Rating derived under the traditional methodology. The adjusted measure scores would be derived from regression models of beneficiary-level measure scores that adjust for the average within-contract difference in measure scores by LIS/DE and/or disability status for MA or PDP contracts without masking potential differences in quality across contracts. This approach is equivalent to case-mix adjustment or patient-mix adjustment in a patient-level linear regression model with contract intercepts and beneficiary-level indicators of LIS/DE/disability status, similar to the approach currently used to adjust CAHPS patient experience measures. Measure scores are adjusted first and then the adjusted measure score is converted to a measure-level Star Rating using the measure thresholds for the given Star Ratings year.

The Categorical Adjustment Index is derived via four steps: (1) contracts are divided into an initial set of categories based on some combination of LIS/DE/disability (this might be 10 categories corresponding to the 10 deciles of LIS/DE or the 16 combinations of LIS/DE quartile and disability quartile, or some other grouping); (2) the mean difference between the Adjusted Overall or Summary Star Rating and the Unadjusted Star Rating is computed within each of the initial categories; (3) the mean differences of the initial categories in step (2) are examined and combined into final adjustment groups such that initial categories with similar means form the groups; and (4) the mean star Adjustment is computed within each of the final adjustment groups - this is the Categorical Adjustment Index.

The Star Rating measure's specification is unchanged. The Categorical Adjustment Index is applied external to the specification and is applied at the Overall and/or Summary Star Rating. Each contract within a given final LIS/DE/disability group receives the same adjustment to its Overall and/or Summary Star Ratings. The index would be determined using the current year's ratings. For the 2017 Star Ratings (measurement year 2015), the Categorical Adjustment Index values would be based on the observed values for the 2017 Star Ratings year using data from contracts that meet current reporting requirements.

The table below depicts an example of the overall method and one possible scenario for groupings employed to determine the Categorical Adjustment Index. In preliminary work, these eight possible final categories of LIS/DE/disability were derived by combining 16 initial combinations of LIS/DE and disability quartiles in the manner described above.

Grouping of contract, based on % LIS/DE and % Disabled	Mean Overall Unadjusted Star Rating	Mean Adjusted Star Rating	Mean Difference in Star Rating (Adjusted - Unadjusted)
LIS/DE 1st quartile & disability 1st quartile			
LIS/DE 1st quartile & disability 2nd-4th quartiles			
LIS/DE 2nd quartile & disability 1st quartile			
LIS/DE 2nd quartile & disability 2nd-4th quartile			
LIS/DE 3rd quartile & disability 1st-2nd quartiles			
LIS/DE 3rd quartile & disability 3rd-4th quartiles			
LIS/DE 4th quartile & disability 1st-3rd quartiles			
LIS/DE 4th quartile & disability 4th quartile			

Indirect Standardization

Indirect standardization, an alternative proposal, would be applied to a subset of the individual Star Ratings measure scores; measure stars are not used because that would incorrectly assume assignments of measure stars are linear in the underlying measure and thus, lead to measurement error. The focus of the adjustment is the within-contract LIS/DE and/or disability status difference while allowing for the existence of true differences in quality by contract. The standardization would employ the current year's ratings.

An expected measure score would be calculated using the percent of LIS/DE and/or disabled and non-LIS/DE and/or non-disabled beneficiaries per measure multiplied by the adjusted mean national performance for each subgroup. As above, this method could use a variety of grouping methodologies to determine appropriate subgroups. These contract-adjusted LIS/DE and/or disabled and non-LIS/DE and/or non-disabled national performance means would be calculated such that they differed by the national performance mean within-contract LIS/DE versus non-LIS/DE difference (and/or disabled versus non-disabled difference). The expected measure score is a weighted average based on the composition of the enrollees of the contract and the national adjusted mean measure score for the subgroups of interest (e.g., LIS/DE and/or disabled versus non-LIS/DE and/or non-disabled beneficiaries). For example: Using indirect standardization at the measure score level, the expected measure score for diabetes control would use the adjusted national means for performance on diabetes control for the subgroups of interest (LIS/DE and non-LIS/DE and/or disabled and non-disabled beneficiaries). To simplify the example, the focus of the standardization is disability status, but the same method could be expanded to more than two subgroups and include standardization for LIS/DE and non-LIS/DE. If the adjusted national mean performance on diabetes control is 50% for disabled beneficiaries and 60% for non-disabled beneficiaries, a contract that served only disabled beneficiaries would be expected to perform at 50% and a contract that served only non-disabled beneficiaries would be expected to perform at 60%. Contracts that had both disabled and non-disabled beneficiaries would be expected to perform at a specific level

between 50% and 60%, as a weighted average based on the composition of the enrollees in the contract for that measure. Each contract would be judged against its expected performance.

Next, the ratio of the observed-to-expected measure score would be calculated. (The observed value is the measure score based on the rating year's data). The observed-to-expected ratio would equal one for contracts that performed at the level expected given their percentages of LIS/DE and/or disabled beneficiaries and indicates average performance. In contrast, a ratio less than one would indicate lower observed performance than expected given the contract percentage of LIS/DE and/or disabled beneficiaries; similarly, ratios greater than one would indicate better than expected performance.

The adjusted measure score would then be calculated. The adjusted measure score is the product of the observed-to-expected ratio for a contract and the adjusted national mean performance for all Medicare beneficiaries. Two contracts with identical observed measure performance but different expected measure performance would receive different adjusted measure scores. The adjusted measure score would be converted to a measure-level Star Rating using the measure thresholds for the given Star Ratings year. The adjusted measure level stars would then be used in the determination of the Overall and/or Summary Star Rating.

The tables below summarize for a single measure the information needed and the overall methodology to determine the adjusted measure score. The first table provides the national values needed to indirectly standardize a measure score. The second table and corresponding information below the table provide a high-level overview of the process of indirect standardization to determine the adjusted measure score. For ease of presentation, the tables focus on an adjustment for the subgroups of LIS/DE and non-LIS/DE, but the same method would be applied for the distinct groupings of LIS/DE, non-LIS/DE, disabled, and non-disabled combinations.

Measure-specific National Adjusted Values	**Adjusted Pass Rate for LIS/DE Beneficiaries**	**Adjusted Pass Rate for non-LIS/DE Beneficiaries**	**Adjusted Overall National Mean**
	A	B	C

Contract	**Proportion of LIS/DE Beneficiaries**	**Proportion of non-LIS/DE Beneficiaries**	**Overall Pass Rate (Observed)**	**Expected Rate (Based on composition of plan)**	**Ratio of Observed to Expected**	**Adjusted Measure Score**
1	D	E	F	G	H	I

The contract's proportion of LIS/DE and non-LIS/DE per measure (values D and E) would be multiplied by the corresponding national adjusted rates for LIS/DE and non-LIS/DE (values A and B) to determine the expected performance for the measure score (value G). The formula applied would be as follows: $\text{Expected Rate} = A \times D + B \times E$.

Next, the ratio of the observed (actual) would be calculated ($\text{Ratio} = \text{F}/\text{H}$). The ratio calculated would be multiplied by the adjusted overall national mean (value C) to determine the adjusted measure score (value I). The adjusted measure score would be converted to a measure-level Star Rating, and the traditional Star Ratings methodology would ensue to determine the adjusted Overall and/or Summary Star Rating for the year.

Additional response to address lack of an LIS indicator for enrollees in Puerto Rico

Notably, Puerto Rican beneficiaries are not eligible for LIS, which is an important element of both potential methodologies. (Beneficiaries in the 50 states are eligible for LIS in the mainland if their income is less than 150% of the Federal Poverty Level). To make the proposed analytical adjustments equitable, CMS is considering an additional adjustment for contracts in Puerto Rico to identify beneficiaries in Puerto Rico's contracts whose incomes would result in an LIS designation in the mainland.

The contract-level modified LIS/DE proportion for Puerto Rico would be developed from two sources of information: (1) the overall proportion of beneficiaries in Puerto Rico with incomes less than 150% of the FPL and (2) each contract's proportion of DE beneficiaries. A linear regression model would be developed to predict the percentage of LIS in a contract using the percentage of DE using all MA contracts except those in Puerto Rico. Preliminary evidence suggests this model has very high accuracy in predicting contract-level LIS from contract-level DE in the 50 states and the District of Columbia, even when restricted to lower-income subsets of states. This model would then be adjusted for use in Puerto Rican contracts (i.e., contracts with a service area only in Puerto Rico) using Puerto Rico's mean percentage of DE and mean of LIS (using the percentage of Puerto Rico's population with incomes less than 150% of the FPL). Using the model developed, each contract's proportion of DE beneficiaries in Puerto Rico would have a corresponding proportion of LIS to create a contract-level measure of LIS/DE percentage to be used in the Categorical Adjustment Index or Indirect Standardization adjustments. We welcome comments on this proposed approach to approximate the percentage of LIS by contract in Puerto Rico, or other possible suggestions of ways to estimate this percentage.

We also considered options to address the unique challenges that Puerto Rican contracts face in improving medication adherence. It has been shown that beneficiaries' out-of-pocket costs may adversely affect medication adherence, which presents an additional barrier for Puerto Rican contracts serving beneficiaries whose incomes would result in an LIS designation in the mainland. One option we considered was to reduce the weights of the three Part D Medication Adherence measures for Puerto Rican contracts. A prior proposal in the 2015 draft Call Letter to reduce the weight of the three Part D Medication Adherence measures to 1.5 as access measures for all Part D sponsors was not supported by the majority of commenters. MA plans and PDP sponsors expressed concerns that this type of change would be contrary to efforts to encourage coordination of care, as well as decrease performance in other quality measures. We simulated this change to assess the potential impact. The highest ratings for substantially all of the Puerto Rican contracts remained unchanged based on simulation of the 2015 Star Ratings data. We commend the Puerto Rican contracts on their improved performance overall across the 2016 Star Ratings and the Part D Medication Adherence

measures in particular. We found no changes to the highest ratings for Puerto Rican contracts when we simulated reducing the weight of the adherence measures using 2016 Star Ratings data. We welcome other proposals to account for the barriers for Puerto Rican contracts to improve medication adherence.

Summary

The potential interim analytical adjustments to address the LIS/DE/disability effect, the Categorical Adjustment Index and Indirect Standardization, share common features. These adjustments were developed while focusing on measuring within-contract differences, while allowing for the existence of true differences in quality by contract. In addition, both methods rely on an adjustment external to the measure specification. The methods result in an adjustment to the measure scores for a subset of measures. The adjusted measure scores are converted to a measure-level Star Rating employing the original measure thresholds for the Star Ratings year before adjustment. For the improvement measure we would use unadjusted scores.

Both analytical adjustments under consideration would require additional processing of the data by CMS. Both methods are applied to the measure score and employ a measure-specific proportion of LIS/DE/disabled which is not constant across measures given measure specification exclusions. Further, both adjustments are based on the current year's data and thus, values needed for the application of the methods (Categorical Adjustment Index values and the adjusted overall national mean per measure and for each subgroup for Indirect Standardization) would not be available until after the submission of the data for the year. Given the additional data processing steps, the analytical adjustments may result in a compressed timeframe for Part C and D contracts' review process of the ratings. The Categorical Adjustment Index values would be incorporated in the Star Ratings Technical Specifications and would be a set of adjustment factors to the overall and/or summary ratings. For the Indirect Standardization method, contracts would need to review a larger volume of intermediate calculations for each measure during the preview periods.

Both of these methods adjust for within-contract differences. Based on our research, we would not expect either approach to generate major adjustments in the overall Star Ratings, though any adjustment may be significant for individual contracts. We recognize adjustments that account for the full between-contract differences (e.g., summary level indirect standardization) would make a considerably larger difference overall, but these approaches could risk over-crediting poor quality contracts. Implementing such an approach coupled with adjustments both upward (for contracts with a high proportion of duals) and downward (for contracts with a lesser proportion of duals) would likely result in larger swings both positively and negatively for contracts. CMS is interested in understanding comments about this approach, including its impact on accurate measurement of quality. CMS also notes that ASPE will continue to explore options that could be implemented in future years.

E. 2017 CMS Display Measures

Display measures on 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 being tested before inclusion into the Star Ratings, or measures displayed for informational purposes. Similar to the 2016 display page, organizations and sponsors have the opportunity to preview their data for the display measures prior to release on CMS' website. Data for 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 2016 display measures will continue to be shown on www.cms.gov. CMS will continue to provide advance notice regarding measures considered for implementation as future Star Ratings measures. Other display measures may be provided as information only.

1. **Timely Receipt of Case Files for Appeals (Part D) & Timely Effectuation of Appeals (Part D).** For the 2016 display measures, the data time frame for both measures was 01/01/2015 – 06/30/2015. CMS proposes to change the data time frame from the first six months of the current year to January 1 – December 31 of the previous year. For example, the 2017 display measures would be based on IRE data from January 1, 2015-December 31, 2015. This change will allow the appeal display measures to match the same timeframe used for the Part D Appeal Star Ratings measures.

The following are a number of new measures for the 2017 display page.

2. **Medication Reconciliation Post Discharge (Part C)**. 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 made two changes: 1) expanded the coverage on this measure from Medicare Special Needs Plans only to all of MA; and 2) expanded the age range to members 18 years and older. Both of these 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. CMS is planning to include this measure on the 2017 display page and is planning to include it in the 2018 Star Ratings.

3. **Hospitalizations for Potentially Preventable Complications (Part C).** NCQA added to HEDIS 2016 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 assesses 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—including coordination of that care—to prevent the complications of chronic and acute conditions that result in hospitalization. CMS is planning to include this measure on the 2017 display page and is planning to include it in the 2018 Star Ratings.

4. **Statin Therapy for Patients with Cardiovascular Disease (Part C).** NCQA has added two sets of statin therapy measures to HEDIS aligned with the 2013 ACC/AHA blood cholesterol guidelines. These 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. Since some of these HEDIS measures overlap with the measures developed by the PQA, CMS is planning to include only one of the HEDIS measures on the 2017 display page and is planning to include this measure in the 2018 Star Ratings. This measure focuses on statin therapy for patients with cardiovascular disease. It is the percentage of males 21 to 75 years of age and females 40 to 75 years of age who were identified as having clinical atherosclerotic cardiovascular disease and were dispensed at least one high or moderate-intensity statin medication during the measurement year.

5. **Asthma Measures (Part C).** NCQA has expanded their asthma measures to include older adults. HEDIS 2016 includes two measures for older adults. Medication Management for People with Asthma is the percentage of members 5 to 85 years of age 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). The Asthma Medication Ratio is 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. CMS is planning to include these on the 2017 display page and will consider these for inclusion in Star Ratings for future years.

6. **Statin Use in Persons with Diabetes (SUPD) (Part D).** This new PQA-endorsed measure, Statin Use in Persons with Diabetes (SUPD), calculates the percentage of patients between 40 and 75 years old who received at least two diabetes medication fills and also received a statin medication during the measurement period. Beneficiaries in hospice according to the Enrollment Database (EDB) will be excluded from the denominator of the SUPD measure for the entire year. Part D sponsors have received year of service 2015 SUPD measure reports on a monthly basis through the Patient Safety Analysis website, and we will add the SUPD measure to the 2017 display page (using 2015 data). We propose adding the SUPD measure to the 2018 Star Ratings (using 2016 data).

Forecasting to 2018 and Beyond

The following describes changes to existing measures and potential new measures. CMS will also monitor any additional measures developed by NCQA or PQA for potential incorporation into the Star Ratings.

F. New Measures:

See section E above which describes a number of new measures under consideration for the 2018 Star Ratings that will be reported as 2017 display measures. The following are additional measures under consideration for the Star Ratings or display measures for 2018 and beyond.

1. **Care Coordination Measures (Part C).** Effective care coordination contributes to improved health outcomes. CMS believes that 5-star contracts 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 well as on data we collect from CAHPS surveys which reflect enrollees' experiences with the care they receive.

CMS is working to expand efforts in this area. To identify potential new care coordination measures, CMS is utilizing experts to conduct targeted research, extensive literature reviews, and data analysis, and to engage in discussions with expert panels and high performing plans. As part of this effort, we are considering various data sources; whether the measures should be focused on subgroups of MA enrollees or all MA enrollees; the activities that best represent care coordination such as ensuring seamless transitions across settings, appropriate follow up after inpatient and emergency department visits, communication across providers, and comprehensive assessments; and the relationship between the plan and provider in care coordination activities. NCQA, using administrative and medical record data, will begin testing the following proposed measures using 2015 data: primary care provider (PCP) notification of inpatient admissions, summary of care record in PCP chart, follow-up with PCP/specialist following hospital discharge or emergency department visit, and in the ambulatory setting whether there is a comprehensive assessment performed and documented by the PCP/specialist and whether there is a specialist visit summary in the PCP chart. Additionally, CMS has recently awarded another contract to develop care coordination measures using administrative data, including MA encounter data and Part D data. CMS welcomes comments on measures that could be developed using MA encounter data. We will continue to provide updates to the industry as this work progresses. Measures developed and tested may be considered for future inclusion on the display page and in Star Ratings.

2. **Depression Measures (Part C).** NCQA has adapted a provider-level depression outcome measure developed by Minnesota Community Measurement for use in HEDIS. Depression Remission or Response in Adolescents and Adults (DRR) uses a patient-reported outcome measure, the PHQ-9 tool, to assess whether patients with depression have achieved remission or have an improvement in their symptoms. The measure assesses the percentage of individuals age 12 and older with depression and an elevated PHQ-9 score (greater than 9) who achieve a PHQ-9 score of less than 5 at six months or have a 50% reduction in their PHQ-9 score. This measure also uses a new data collection methodology for HEDIS, relying on data coming from electronic clinical data systems (e.g., EHRs, clinical registries, case management records). If approved, the new measure would be published in HEDIS 2017.

3. **Appropriate Pain Management (Part C).** NCQA is exploring opportunities to develop a new measure(s) focusing on appropriate pain management. The intent is to assess the quality of pain management and treatment. There is no definite timeline established for the development of this measure.

4. **Use of Opioids from Multiple Providers or at High Dosage in Persons without Cancer (Part D).** In the 2016 Call Letter, we noted that three opioid overutilization measures were in development by the PQA. We further stated that if these measures were endorsed by the PQA prior to the 2017 bid deadline in June 2016 that we may adopt them as future display measures or alternatively use them in the Overutilization Monitoring System (OMS). The measures were endorsed by the PQA in May 2015.

   PQA's three opioid measures examine multi-provider, high dosage opioid use among individuals 18 years and older without cancer and not in hospice care.

Measure 1 (Opioid High Dosage): The proportion (XX out of 1,000) of individuals without cancer or hospice receiving prescriptions for opioids with a daily dosage 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 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 hospice receiving prescriptions for opioids with a daily dosage 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.

We tested the measures using the PQA specifications. We will develop new patient safety opioid overutilization measure reports (beginning with 2016 dates of service) to provide to Part D sponsors on a monthly basis through the Patient Safety Analysis website, similar to the other patient safety measures. The website also includes the OMS. The reports will allow sponsors to track their performance over time and allow for contract level trending and outlier analyses. We will also add these three measures to the 2018 Part D display page (using 2016 data). We do not recommend adding these measures to the Star Ratings at this time due to concerns (1) about the current lack of consensus clinical guidelines for the use of opioids to treat chronic, non-cancer pain and potential exceptions due to medical necessity and (2) pending additional analysis on diagnosis data sources, such as newly available encounter data for Medicare Part C and resolving timing issues of RAPS file updates, which are used to identify exclusions for certain cancer conditions.

Additionally, NCQA is adapting the three opioid overuse measures developed by the PQA for potential use in HEDIS.

5. **Antipsychotic Use in Persons with Dementia (APD) (Part D).** CMS has been particularly concerned with the unnecessary use of antipsychotic drugs in nursing homes and, as a result, has pursued strategies to increase awareness of antipsychotic use in long term care settings. In 2013, we began to calculate a general atypical antipsychotic utilization rate, called *Rate of Chronic Use of Atypical Antipsychotics by Elderly Beneficiaries in Nursing Homes*, for inclusion in the Part D display measures. The average rates decreased from approximately 24.0% in 2011 to 21.4% in 2013.

   There continues to be increased attention on this important issue. The United States Government Accountability Office (GAO) released a report[4] in January 2015 describing the inappropriate use of antipsychotics in Part D beneficiaries with dementia, in both community (i.e., outside of nursing homes) and long-stay nursing home residents during 2012, with

[4] Antipsychotic Drug Use: HHS Has Initiatives to Reduce Use among Older Adults in Nursing Homes, but Should Expand Efforts to Other Settings. http://www.gao.gov/products/*GAO-15-211*. *GAO-15-211*: Published: Jan 30, 2015. Publicly Released: March 2, 2015

recommendations for CMS to address this problem. The GAO conducted this study due to concerns raised regarding the use of antipsychotic drugs to address the behavioral symptoms associated with dementia, the FDA's boxed warning that these drugs may cause an increased risk of death when used by older adults with dementia, and because the drugs are not approved for this use.

In addition, the PQA endorsed the measure, *Antipsychotic Use in Persons with Dementia (APD).* This provides CMS with a new measure developed through a consensus process to monitor the inappropriate use of antipsychotics in both the nursing home and community settings across Medicare Part D plans.

We tested this measure based on the PQA specifications. The APD measure rate was calculated for all contracts, MA-PDs, PDPs, and at the individual contract-level for all beneficiaries, community-only residents (never a nursing home resident), and both short-term and long-term nursing home residents that met the inclusion and exclusion criteria. Beneficiaries were identified as long-stay nursing home residents if they had stays greater than 100 cumulative days in a nursing home during the year based data in the Long Term Care Minimum Data Set (MDS). Each beneficiary was counted in only one category for the entire measurement period within a contract, not considered separately for time spent in different settings (e.g., a beneficiary who experienced both short-term and long-term nursing home stays was included only in the long-term category).

To identify the numerator and denominator populations, we used diagnosis data obtained from inpatient (IP), outpatient (OP), and carrier claims from the Common Working File (CWF) and RxHCCs from the RAPS. OP and Carrier claims are available for PDP contracts only. We also adjusted rates based on the number of months beneficiaries are enrolled in each Part D contract (i.e., member-years adjustment).

We conducted reliability testing using mixed effect logistic regression with varying intercept. The testing results indicate that the rate variations at the contract level are statistically significant, providing evidence that the measure is reliable.

We will develop new patient safety APD measure reports to provide to Part D sponsors on a monthly basis through the Patient Safety Analysis website beginning with year of service 2016. We also recommend adding the overall APD measure plus breakout rates for community-only residents, short-term nursing home residents, and long-term nursing home stay residents to the 2018 Part D display measure set (using 2016 data) to continue to draw attention to the inappropriate use of antipsychotics in persons with dementia without an appropriate mental health diagnosis in both the community and nursing home settings. The APD measure will replace the *Rate of Chronic Use of Atypical Antipsychotics by Elderly Beneficiaries in Nursing Homes* display measure. However, we do not recommend adding this measure to the Star Ratings pending additional research on diagnosis data sources, such as newly available encounter data for Medicare Part C and resolving timing issues of RAPS file updates.

G. Changes to Existing Star Ratings and Display Measures and Potential Future Changes:

1. **Colorectal Cancer Screening (Part C Star Rating)**. The Colorectal Cancer Screening (COL) measure assesses the percentage of adults age 50-75 years of age who had appropriate screening for colorectal cancer. This measure is based on the U.S. Preventative Services Task Force (USPSTF) guideline on colorectal cancer screening in adults age 50-75. NCQA is monitoring updates to the guideline as the USPSTF has released a draft recommendation statement for public comment. We have discussed the guideline timing with the Agency for Healthcare Research and Quality team in charge of the USPSTF process, and they note the final release is not likely to occur until late 2016. NCQA will consider revisions to the COL measure once the USPSTF final recommendation statement is published.

2. **Fall Risk Management (Part C Star Rating).** The Fall Risk Management (FRM) measure, collected through the Health Outcomes Survey, consists of the following two indicators: 1) *Discussing Fall Risk* assesses the percentage of Medicare members 75 years of age and older or 65-74 years of age with a balance or walking problem or fall in the past 12 months who discussed falls or problems with balance or walking with their current practitioner; and 2) *Managing Fall Risk* assesses the percentage of Medicare members 65 years of age and older who had a fall or had problems with balance or walking in the past 12 months and received fall risk intervention from their current practitioner defined as suggesting use of a cane or walker, a vision or hearing test, physical therapy or exercise, or taking of a postural blood pressure. NCQA is currently re-evaluating this measure to align with the most current U.S. Preventive Services Task Force (USPSTF) guidelines. NCQA is proposing to 1) revise the denominator in the Discussing Fall Risk indicator to include all Medicare members age 65 and older and 2) revise the numerator for the Managing Fall Risk indicator to include use of vitamin D and remove vision or hearing test and taking of postural blood pressure. These proposed changes, if approved, would be published in HEDIS 2017 or HEDIS 2018.

3. **Pneumococcal Vaccination Status for Older Adults (Part C Display).** The Pneumococcal Vaccination Status for Older Adults (PNU) measure, collected through the Medicare CAHPS survey, assesses the percentage of Medicare members 65 years of age and older who have ever received a pneumococcal vaccination. In 2014, The Advisory Committee on Immunization Practices (ACIP) released new recommendations that all adults 65 years of age and older should receive sequential administration of both PCV13 and PPSV23. NCQA is considering changes to the measure to align with the most current guidelines. There is no definite timeline established for these changes. This measure is on the CMS display page.

4. **CAHPS measures (Part C & D).** The current MA & PDP CAHPS Survey includes the core CAHPS 4.0 Health Plan Survey. CMS conducted an experiment in 2015 to understand how CAHPS measures differ between 4.0 and 5.0, and based on the results we propose to update the survey for future years to reflect AHRQ's CAHPS 5.0 Health Plan Survey. The findings from the experiment suggest that these changes are associated with a small increase in scores for several evaluative MA measures. These small increases did not significantly differ across contracts. Since there are no longer fixed thresholds for Star Ratings and they are based on the actual distribution of scores, there should be no shifts in Star Ratings due to transition to the version 5.0 instrument compared to what would have been the case with 4.0. Every

contract would have the same expected Star Rating whether version 4.0 or 5.0 is used, and the correlation between this year's Star Ratings and next year's Star Ratings should be the same regardless of whether 4.0 or 5.0 is used next year.

The 5.0 update applies recent improvements in survey design that resulted from development and testing of the Clinician & Group Surveys. The 5.0 version of the CAHPS Health Plan Survey incorporates some minor changes into the wording of core items, and a change in the placement of one core item that also resulted in the deletion of a screener item. The following are the changes in the 5.0 version of the Health Plan Surveys:

- **The items about access to urgent and non-urgent appointment items** were modified to ask respondents if they were able to get an appointment as soon as they needed, as opposed to as soon as they *thought* they needed. Non-urgent appointments are described as *a check-up or routine care* rather than *health care.* In addition, the phrase*, "…not counting the times you needed care right away"* was deleted from these questions*.* These revisions simplify the items and make them consistent with questions in other CAHPS surveys.

- **The item about how often it was easy to get appointments with specialists** was revised to ask respondents if they got an appointment to see a specialist as soon as they needed. This revision makes the item consistent with other CAHPS items that ask about access to care.

- **The item about how often it was easy to get care, tests, or treatment** was moved from the Your Health Plan section to the Your Health Care in the Last 6 Months section, because respondents had difficulty attributing this item to the health plan.

- **The screener item about getting care, tests, or treatment through the health plan** was deleted because the subsequent question was moved to an earlier section of the survey and no longer required a screener.

These changes would take effect for the 2017 CAHPS survey administration (2018 Star Ratings) based on OMB approval. Since we are modifying question wording, we propose the following standard for deciding that a specification change has occurred for a CAHPS measure for the purposes of excluding it from the improvement measure calculation: (1) at least one item within the measure changed in wording, had a wording change in its screener, or had a wording change in the immediately preceding item, and (2) the measure score in version 5.0 was significantly different from the measure score in version 4.0 in the 5.0 experiment. Three MA measures met this standard: Getting Care Quickly, Customer Service, and Care Coordination; thus, these three measures would be excluded from the Part C improvement measure for the 2018 Star Ratings.

We are also considering changing the sampling for CAHPS when a contract is listed in HPMS as a consolidation, merger, or novation between July of the prior year and January of the current year when the CAHPS sample is drawn. We are considering changing the sampling frame for

the surviving contract to include the enrollees for all members of all contracts involved if the two or more contracts merging, consolidating or novating are under the same parent organization. This would go into effect for the 2017 sample draw.

5. **Medication Adherence for Hypertension (RAS Antagonists) (Part D Star Rating).** Based on PQA specification change, the measure will exclude from the denominator those patients with one or more claims for sacubitril/valsartan.

6. **MPF Price Accuracy (Part D Star Rating).** As stated in the 2016 Call Letter, CMS is considering a few updates to this measure for the 2018 Star Ratings. The first proposed change is related to the method by 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 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 retail pharmacies that are also long term care, mail order, or home infusion pharmacies 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 60-62 and 90-93 day supplies. 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 day supply Pricing File data, to compare MPF and PDE costs. While the majority of claims are for a 30 day supply, we found that claims with a 90 day 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. CMS began requiring pharmacies to populate the Pharmacy Service Type field on all PDEs at the end of February 2013. We recommend expanding the retail claims identification process to include all PDEs that are from retail pharmacies according to the Pharmacy Cost data and have a Pharmacy Service Type of either Community/Retail or Managed Care Organization (MCO). Although some sponsors cited concern about the accuracy of these data as reported by pharmacists, Part D sponsors are ultimately responsible for the accuracy of their submitted PDE to CMS. According to PDE requirements, CMS expects "…sponsors and their network pharmacies to develop and implement controls to improve the accuracy of this information during 2013…" This methodology change would increase the number of PDEs eligible for inclusion in the Price Accuracy Scores while continuing to identify only retail 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. CMS is aware that while the MPF display is updated every two weeks, real time pricing, at the point of sale, can change as often as every day. Some sponsors have expressed concern that in order to perform well in the Price Accuracy measure, they cannot offer lower prices at point of sale in real time than the prices are displayed on MPF. We would note that PDEs priced lower than MPF displayed pricing do not lower a contract's score in this measure. CMS' simulation of this proposal found little change in the range of contracts' accuracy scores. Other options we explored include 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.

As noted in the 2016 Call Letter, we propose that these changes are implemented for the 2018 Star Ratings (using 2016 PDE and MPF data). 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.

7. **Drug-Drug Interactions (DDI) (Part D Display).** The PQA-endorsed DDI measure is currently a Part D display measure. This measure is defined as the percent of Medicare Part D beneficiaries who received a prescription for a target medication during the measurement period and who were dispensed a prescription for a contraindicated medication with or subsequent to the initial prescription.

The PQA has conducted an extensive review of the drug-drug pairs included in the DDI measure. They engaged a DDI expert panel convened by the University of Arizona on PQA's behalf, which completed the review, including a comparison to the DDI list developed for the Office of the National Coordinator for Health Information Technology (ONC). Next, the Expert Panel's recommendations will be reviewed by the PQA's Measure Update Panel for consideration by the PQA's Quality Metrics Expert Panel (QMEP). We anticipate that there will be extensive changes to the DDI measure specifications. We will closely monitor any updates to this measure, test updated specification when available, and propose changes in the future for the Part D display measure and patient safety reporting.

8. **Center for Medicare and Medicaid Innovation Model Tests.** The CMS Center for Medicare and Medicaid Innovation has announced the Medicare Advantage Value-Based Insurance Design (MA-VBID) and the Part D Enhanced MTM model tests. Beginning January 1, 2017, in a

limited number of states, CMS will give MA only, MA-PD or Part D plans participating in these tests additional flexibilities intended to improve the quality of care and reduce costs in the Medicare Advantage or Part D programs, respectively. More information about the specific flexibilities offered in these model tests is available at https://innovation.cms.gov/initiatives/HPI.

Some stakeholders have expressed to CMS the potential for the improvements in quality in these models to favorably influence the Star Ratings of contracts with participating plans, as compared to the performance of those ineligible to participate. The goal is to not penalize participants or non-participants. As the model tests are implemented, we will closely monitor performance trends of participating plans across individual measures and determine if any changes are warranted. We welcome any comments on how to address any potential differences in performance between participating and non-participating plans.

The Part D plans participating in the Part D Enhanced MTM model test will be waived from the MTM requirements under Section 1860D–4(c)(2) and 42 CFR 423.153(d) and the Part D reporting requirements for MTM. However, Part D sponsors will not be waived from establishing MTM programs in compliance with current requirements and reporting data for the remaining plans under each Part D contract. Therefore, the MTM Program CMR Completion Rates will be calculated using available plan-reported data from the remaining plans under the Part D contract.

H. Measurement and Methodological Enhancements

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 issues not described in earlier sections.

Appendix

Improvement measures (Part C & D):

Part C or D	Measure	Measure Type	Weight	Improvement Measure
C	Breast Cancer Screening	Process Measure	1	Yes
C	Colorectal Cancer Screening	Process Measure	1	Yes
C	Annual Flu Vaccine	Process Measure	1	Yes
C	Improving or Maintaining Physical Health	Outcome Measure	3	No
C	Improving or Maintaining Mental Health	Outcome Measure	3	No
C	Monitoring Physical Activity	Process Measure	1	Yes
C	Adult BMI Assessment	Process Measure	1	Yes
C	Special Needs Plan (SNP) Care Management	Process Measure	1	Yes
C	Care for Older Adults – Medication Review	Process Measure	1	Yes
C	Care for Older Adults – Functional Status Assessment	Process Measure	1	Yes
C	Care for Older Adults – Pain Assessment	Process Measure	1	Yes
C	Osteoporosis Management in Women who had a Fracture	Process Measure	1	Yes
C	Diabetes Care – Eye Exam	Process Measure	1	Yes
C	Diabetes Care – Kidney Disease Monitoring	Process Measure	1	Yes
C	Diabetes Care – Blood Sugar Controlled	Intermediate Outcome Measure	3	Yes
C	Controlling Blood Pressure	Intermediate Outcome Measure	3	Yes
C	Rheumatoid Arthritis Management	Process Measure	1	Yes
C	Reducing the Risk of Falling	Process Measure	1	Yes
C	Plan All-Cause Readmissions	Outcome Measure	3	Yes
C	Getting Needed Care	Patients' Experience and Complaints Measure	1.5	Yes
C	Getting Appointments and Care Quickly	Patients' Experience and Complaints Measure	1.5	Yes
C	Customer Service	Patients' Experience and Complaints Measure	1.5	Yes
C	Rating of Health Care Quality	Patients' Experience and Complaints Measure	1.5	Yes
C	Rating of Health Plan	Patients' Experience and Complaints Measure	1.5	Yes
C	Care Coordination	Patients' Experience and Complaints Measure	1.5	Yes
C	Complaints about the Health Plan	Patients' Experience and Complaints Measure	1.5	Yes
C	Members Choosing to Leave the Plan	Patients' Experience and Complaints Measure	1.5	Yes
C	Beneficiary Access and Performance Problems	Measures Capturing Access	1.5	No
C	Health Plan Quality Improvement	Improvement Measure	5	No
C	Plan Makes Timely Decisions about Appeals	Measures Capturing Access	1.5	Yes
C	Reviewing Appeals Decisions	Measures Capturing Access	1.5	Yes
C	Call Center – Foreign Language Interpreter and TTY Availability	Measures Capturing Access	1.5	Yes
D	Call Center – Foreign Language Interpreter and TTY Availability	Measures Capturing Access	1.5	Yes
D	Appeals Auto–Forward	Measures Capturing Access	1.5	Yes
D	Appeals Upheld	Measures Capturing Access	1.5	Yes
D	Complaints about the Drug Plan	Patients' Experience and Complaints Measure	1.5	Yes
D	Members Choosing to Leave the Plan	Patients' Experience and Complaints Measure	1.5	Yes
D	Beneficiary Access and Performance Problems	Measures Capturing Access	1.5	No
D	Drug Plan Quality Improvement	Improvement Measure	5	No
D	Rating of Drug Plan	Patients' Experience and Complaints Measure	1.5	Yes
D	Getting Needed Prescription Drugs	Patients' Experience and Complaints Measure	1.5	Yes
D	MPF Price Accuracy	Process Measure	1	No
D	Medication Adherence for Diabetes Medications	Intermediate Outcome Measure	3	Yes
D	Medication Adherence for Hypertension (RAS antagonists)	Intermediate Outcome Measure	3	Yes
D	Medication Adherence for Cholesterol (Statins)	Intermediate Outcome Measure	3	Yes
D	MTM Program Completion Rate for CMR	Process Measure	1	Yes