Summary of Comments to the Request for Comments on 2017 Star Ratings and Beyond

On November 12, 2015, CMS released a memo, Request for Comments: Enhancements to the Star Ratings for 2017 and Beyond, to Part C and D sponsors, stakeholders and advocates. The memo described CMS' proposed methodology for the 2017 Star Ratings for Medicare Advantage (MA) and Prescription Drug Plans (PDP). We received approximately 90 comments representing plan sponsors, associations, consumer groups, and measurement development organizations. Concerns about and requests for clarification of specifications have been passed along to measure developers and stewards. This document provides a summary of the comments received and how we addressed these comments in the draft 2017 Call Letter.

A. *Changes to Measures for 2017*

1. **Improvement measures (Part C & D).**

   Summary of Comments: Our proposal to use the 2014 CAHPS measure score (used in 2015 Star Ratings) as the baseline for the 2017 improvement calculation for that measure 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 was not supported by the majority of commenters. The following reasons were cited:

   - The proposal does not provide an accurate reflection of more recent efforts that may have been made to impact CAHPS performance.

   - Sponsors may be disadvantaged by use of dated 2014 CAHPS measure scores.

   These commenters were either against using prior year CAHPS data or proposed to exclude the measure from the improvement measure. About half requested additional clarification.

   A small number of additional comments related to the improvement measures included:

   - Excluding the Call Center measure or applying revisions due to transparency in the reporting of plan performance or to increase the metric's sample size or tailor the calls/ make test calls to test availability of interpreters.

   - Including either the Part C or the Part D improvement measure to benefit the rating, as well as reducing the current weight of 5 to 3.

   - Excluding the MTM measure from the improvement measure.

   - Considering achievement levels along with grading on the curve for some measures or receiving the improvement measures above a 3-star rating.

   Response: CMS appreciates these concerns. We will use the 2014 CAHPS data only if there is a significant improvement from 2014 to 2016 when they do not have 2015 data due to very low reliability. This policy would affect very few contracts, but this would hold contracts harmless from missing CAHPS data.

1. **Reviewing Appeals Decisions/Appeals Upheld measures (Part C & D).**

   Summary of Comments: Almost all commenters supported the proposed change to include appeal cases that are reopened and decided prior to May 1, 2016 in the upheld measure. Additionally, some commenters suggested:

   - Extending the time frame to include reopened cases past May 1st.

   - Removing cases where the IRE obtains new or different information in making a decision,

   - Providing sponsors with the same reports that CMS receives.

   - Including reopening cases that occur in the following year's data.

   - Accounting for the volume of cases appealed.

   - Adjusting the threshold for contracts to be excluded based on the contract's membership.

   Response: CMS will move forward with changing the reopening deadline to May 1. CMS will review other suggestions.

2. **Contract Enrollment Data (Part C & D).**

   Summary of Comments: Most commenters supported the proposed change. Some commenters did not understand that we used the enrollment in the Special Needs Plans (SNP) Care of Older Adults (COA) measures only for submissions that do not contain a valid eligible population element.

   Response: Further review by CMS has shown that the proposed change is no longer necessary due to changes in the way enrollment data are processed over time.

3. **Transition from ICD-9 to ICD-10 (Part C & D).**

   Summary of Comments: Almost all commenters suggested the transition to ICD-10 is appropriate. Some suggested errors are likely because of the complexity of ICD-10 and that plans should be held harmless for errors the transition might cause. Many noted NCQA's guidance and asked CMS to provide similar guidance and/or to prompt PQA to provide similar guidance. Some suggested PQA has not anticipated likely challenges, while others worry PQA measures may require reference to diagnostic codes that PDPs and pharmacies would not be able to access.

   Response: PQA measures currently used in Star Ratings do not reference ICD-9 diagnostic codes, so CMS will clarify that the transition to ICD-10 is not relevant to those measures at this time. CMS will encourage measure stewards to provide guidance about the transition, if and as needed.

4. **Appeals Upheld measure (Part D).**

   Summary of Comments: Several commenters disagreed with the proposal to again include cases for beneficiaries enrolled in hospice in this measure.

   - Sponsors were concerned that data from hospice-enrolled beneficiaries are unreliable and not reflective of plan performance.

- Commenters also noted that appeals from members in hospice do not relate to services rendered by the MA; including those in the calculation may result in a rate which does not represents the true compliance of the MA.

- Additionally commenters are concerned that they may be negatively impacted as they may not have the visibility to hospice status at the time of the initial review due to the lag in timing of receipt of hospice indicators and/or retroactive changes to hospice status as sent on the TRR.

A few commenters agreed with the proposed change. One commenter also suggested removing cases where the IRE obtains new or different information in making a decision and to align time frames and processes for plan sponsors and IREs to make a more equitable evaluation of plan sponsor decisions.

Response: As noted in the 2016 Call Letter, this exclusion was only necessary for the 2016 measure as it was based on 2014 data that may have been affected by policy changes in 2014. CMS policy has not changed since 2014, and there is no reason to exclude hospice appeal cases from the 2017 Star Rating Appeals Upheld measure.

5. **Medication Therapy Management (MTM) Program Completion Rate for Comprehensive Medication Reviews (CMR) measure (Part D).**

   Summary of Comments: We received only positive feedback for CMS' inclusion of more detailed data during HPMS plan previews. Sponsors supported CMS program audits of MTM, but many voiced concerns about the specific audit standards to be used, if sponsors could be penalized due to MTM program variations, and how these audits would differ from the current data validation activities for plan-reported data.

   Response: In the Data Integrity section, CMS will clarify that more information about MTM audit criteria will be released soon; questions should be directed by email to part_c_part_d_audit@cms.hhs.gov. We will also clarify that we will not apply any relevant MTM program audit findings that could demonstrate data integrity issues for sponsors that participate in the MTM audit during the pilot. We will also clarify that Data Validation standards assess compliance to CMS' reporting requirements and technical specifications while CMS program audits are more comprehensive assessments of the contracts' MTM programs.

B. *Removal of Measures from Star Ratings*

   1. **Improving Bladder Control (Part C).**

      Summary of Comments: Most commenters expressed some doubt about the validity, reliability or utility of this measure and supported removing the measure from Star Ratings, and giving plans advance notice if the measure were to return to the Star Ratings. Some plans thought expecting improvement in bladder control unrealistic for their (Special Needs) population or not meaningful given the other conditions (e.g., receiving dialysis for kidney failure) their enrollees may have. Many asked for clarification if CMS intends to return the measure to Star Ratings in 2018, which would be the normal course of action. Some stated they would prefer this measure were dropped entirely. Some commenters thought the focus of the measure should move beyond receipt of treatment, while others preferred the focus remains on receipt of treatment.

There was also confusion about whether this measure was a cross-sectional measure or a two-year change score.

Response: CMS will clarify that the HEDIS measures derived from the HOS survey, like Improving Bladder Control, are cross-sectional and do not require comparison of responses from the same cohort two years later. CMS will forward comments to the measure steward, NCQA.

2. **High Risk Medication (Part D).**

Summary of Comments: A majority of the commenters supported removing the High Risk Medication (HRM) measure from Star Ratings and moving to the display measures for 2017. Over one-quarter of the commenters supported removal of the HRM measure from Star Ratings, but for 2018 or beyond, noting that plan sponsors have made significant investments to improve HRM performance. Few commenters did not support this change. A number of commenters recommended sufficient lead time before future updates to the measure specifications are implemented to consider formulary and bid timelines. Some commenters suggested exclusion of hospice patients from the measure calculation.

Response: CMS will move forward with the proposal to remove the HRM measure from the Star Ratings and move to the display measures for 2017. We shared specification comments with the PQA.

The timing of implementation of updates to the measure specifications in the future will provide sufficient lead time ahead of the formulary and bid deadlines.

C. *Data Integrity*

Summary of Comments: Many supported the development of program audits of Part D sponsors' MTM programs, and voiced similar concerns as described in section A6 above. A few commenters complained about CMS policy to reduce measures due to Data Validation failures at an element level, or in general that this policy is unwarranted, and that reductions to 1 star should only be made if sponsors intentionally made data errors. One commenter suggested that if CMS or its contractors made errors, then affected sponsors should be assigned 5 stars, and/or be able to reuse the prior year's data.

Response: As noted in A6's response, we will clarify that information about the MTM audit criteria will be released soon, and that findings identified during pilots of the MTM audit criteria would not be used. We will clarify that reductions are based on systemic failures. A reporting section's overall DV score may be high enough to consider the contract has passed, but specific element-level failures show inaccurate data that would be used in the measure calculation. For example, if the Data Validation found errors in the numbers of beneficiaries enrolled in the MTM, or receiving CMR – but the overall MTM DV score was above 95, CMS would still have concerns about the MTM CMR numerator and denominator.

CMS and its contractors also share in the responsibility for accurate collection and calculation of Star Ratings. Each year, CMS reviews the quality of the data across all measures, variation among organizations and sponsors, and the accuracy and validity of measures before making a final determination about inclusion in the Star Ratings. Rarely, as a result of these reviews, CMS has needed to exclude specific measures from the Star Ratings because systemic errors were found in

the data; these errors may not have been caused by sponsors' processes or actions. While CMS considers all options prior to exclusion of a measure, it would be incorrect to infer contracts' performances from historical data, or reuse prior years' data as individual contract measure performance changes from year to year. This would also unfairly disadvantage contracts, as use of prior year scores may result in lower Star Ratings, versus if the measure was excluded. Using data from a prior collection year for some but not all contracts (i.e., holding contracts harmless for downward adjustment), or assigning 5 stars for all contracts could also appear arbitrary.

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

Summary of Comments: There was widespread appreciation of the attention and careful examination of the LIS/DE/disabled effects that CMS has completed at this time and continues to research. Respondents agreed with the overall approach and focus of the research conducted. Many commenters applauded our efforts and were grateful for the acknowledgement of the issue and the development of the two proposed interim analytical adjustments set forth in the Request for Comments: a Categorical Adjustment Index or Indirect Standardization. Respondents valued the numerous opportunities afforded for stakeholder input. In addition, many commenters supported the engagement of the measure stewards and look to their research as part of the answer to the long-term solution to address any sensitivity of the Star Ratings measures to the enrollment of beneficiaries. Further, several respondents complimented CMS for maintaining the integrity of the core of the Star Ratings Program and the transparent manner in which CMS approached the work related to the issue.

Many respondents mentioned the value of the information provided during the December 3rd User Call, however, some still desired additional clarity of the methodologies. Most respondents requested simulated results so that sponsors could better understand the methodologies employed for the analytical adjustments and the impact of the adjustments on their scores. Further, numerous respondents requested the additional details of the interim adjustments and simulation results in advance of the draft 2017 Call Letter.

Overall, the reaction to the proposed analytical adjustments was mixed. The comments received did indicate a need for further outreach so that stakeholders would be comfortable supporting one of the options proposed to address the LIS/DE/ disabled effect. Many respondents felt they were unable to state a preference of one method over the other until they were able to review simulation results. Of the respondents that did express a preference to a single analytical adjustment, more commenters preferred the Categorical Adjustment Index (CAI) over the method of Indirect Standardization (IS).

There was a request for the release of any research that focused on PDPs. Some commenters expressed a need for details and discussion on the application of the analytical adjustments for PDPs including the availability of the disability status of beneficiaries. Some respondents expressed the need for immediate action and financial relief for plans that served vulnerable populations. Several commenters wanted the interim adjustments to result in meaningful differences in the ratings. Respondents were unsure of the mechanics of the application of the adjustments for plans where beneficiaries were exclusively LIS/DE and others, requested the minimum proportion of beneficiaries in the subgroups of LIS/DE and/or disabled status to qualify for an adjustment. There were inquiries regarding the stability of the adjustments over time, the details of operationalizing the methods, and the strengths and weaknesses of each method.

There were a number of respondents that encouraged CMS to wait rather than to move forward with either of the proposed interim adjustments. Some respondents who preferred a delay referenced either looking to the work of the measure developers or the release of the Assistant Secretary of Planning and Evaluation's (ASPE) report due in the upcoming year as required by the IMPACT Act. One commenter mentioned the need to 'proceed with caution to avoid either creating a double standard of care or lowering standards for chronic disease management.' A respondent questioned the value of an interim adjustment given the increase in the administrative burden for both CMS and sponsors. Another reason cited for waiting for implementation was that the proposed analytical adjustments did not adequately address the issue. In terms of the timing for a policy response, one commenter suggested optional participation in 2017, while another suggested a two-year delay in implementation.

Numerous commenters believed that CMS should further investigate or control for factors beyond LIS/DE and/or disabled status. The attributes that respondents mentioned for additional examination and/or consideration included: age, race, health literacy, household income, homelessness, hunger, transportation challenges, low literacy, unemployment, number of providers, risk scores, medical complexity, chronic conditions, and geographical or regional disparities. Some respondents mentioned that risk adjustment of quality measures would penalize plans that deliver high quality of care irrespective of the population served.

There were a limited number of comments to help guide the selection of the measures to be adjusted. Some commenters mentioned that only measures that revealed a meaningful within-contract difference in our research should be candidates for adjustment. Other commenters believed all Star Ratings measures, regardless of whether they are already adjusted, should be included as candidates for adjustment. A respondent did note that if all measures were adjusted, it would still result in only a small proportion of the measures in the program. There was a comment that stated that any risk adjustment must be rooted in evidence and not disincentivize plans from enrolling vulnerable beneficiaries. Another respondent believed that some risk adjustment done prudently is better than none. One commenter would not support adjusting for medication adherence measures and believed that LIS beneficiaries in their plan do better and thus, the plan would be negatively impacted by such an adjustment. There were several commenters inquiring about the method of reporting – hybrid or census – and its impact on the interim adjustment methodologies.

Some commenters expressed the need for accounting for LIS-look-alikes in the adjustments. Respondents believe that there are beneficiaries that do not quality for LIS but share common characteristics with LIS beneficiaries, such as community factors and income levels, and may in fact experience worse outcomes due to their non-LIS status. The look-alikes therefore impact the within-contract differences due to the modeling that uses a dichotomous variable for LIS status.

Many of the general comments related to the two options CMS is exploring for interim analytical adjustments expressed concern about the possible bias an adjustment may introduce in the Star Ratings program. A respondent expressed the need to capitalize on the heterogeneity of the data. One commenter specifically discussed the importance of maintaining the quality signal that the Star Ratings provide. To compensate for the perceived advantage plans with a high percentage of LIS/DE/disabled enrollees may realize, many commenters supported a hold-harmless provision. (Such a provision would ultimately result in only positive adjustments to the

Overall and Summary Star Ratings.) Several commenters expressed concern regarding the use of the pre-adjusted cut points for the conversion of the adjusted measure score to a measure-level Star Rating and requested justification for this aspect of the methodology. One commenter felt that mixing unadjusted and adjusted measure scores and stars would result in apple to orange comparisons. A respondent asked for simulations using both adjusted and unadjusted cut points.

Numerous respondents believed that the adjustments would increase the complexity of the Star Ratings and impinge on its transparency. In addition, there was concern about the possibility of a shortened plan-preview period. Several commenters suggested an additional plan preview period if CMS were to move forward with one of the interim adjustments. A number of commenters wanted justification of the use of unadjusted measure cut points for the conversion of the adjusted measures score to measure-level Star Ratings. Some commenters were unsure how the adjustment would be applied to plans that were almost exclusively comprised of Dual beneficiaries. Several respondents were unsure when the proposed adjustment would be applied in relationship to the Reward factor. Further, several comments inquired about which measures scores would be used for determining the integration factor – adjusted or unadjusted measures scores.

Many of the comments received that were specific to the CAI were related to the number of initial and final adjustment categories employed in the method.   Commenters were concerned about the collapsing of the initial groups and final adjustment categories and the impact of the groupings on the adjustments and possible misclassification. Respondents who preferred CAI over IS cited reasons such as: easier to understand, similarity to CAHPS, greater transparency, ability to have the adjustment factors in advance of the plan preview, flexibility and accuracy of the method. A number of commenters suggested using additional covariates in the model such as race, gender, and community factors. Some commenters wanted to know more about the stability of the adjustment factors over time.

A number of the comments received that were specific to IS noted the increased complexity of the method. Many commenters expressed concern about the validation needed if the method was employed and the impact on the plan preview period. Some respondents felt the adjustment was more tailored to their specific plan and perhaps, more accurate. Several commenters believe geographic comparisons instead of national comparisons should be used for the standardization process. There were a number of comments regarding the determination of the national means and the impact of its value on measures when the information is calculated based on mixed reporting methods (census or hybrid).

There were a limited number of responses related to the additional response to address lack of an LIS indicator for enrollees in Puerto Rico. Overall, there was appreciation to CMS for addressing the unique circumstances and challenges in Puerto Rico. Respondents expressed concern about the accuracy of the proposed LIS Indicator used for adjustment based on the relationship developed using mainland data and modifying it for use in Puerto Rico. Further, the commenters suggested an adjustment of the medication adherence measures for contracts operating in Puerto Rico.

Response: While the measure stewards are undertaking a comprehensive review of their measures used in the Star Ratings Program and ASPE is continuing its work under the IMPACT Act, CMS is proposing to implement the CAI as an interim analytical adjustment for 2017 Star

Ratings to take into account the impact of LIS/DE and/or disability status on Star Ratings measures. The CAI 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. The proposed interim solution adheres to certain core CMS principles, such as not permitting a lower standard of care for vulnerable beneficiaries, proposing adjustments that reflect the actual magnitude of the differences observed in the data, and recognizing the need for options that are both transparent and feasible for the plans and CMS to implement. The proposal relies on an adjustment external to the measure specifications, as well as, maintaining the integrity of the Star Ratings and the core of its methodology. For contracts operating in Puerto Rico, we plan to proceed with estimating an LIS Indicator, while other data sources continue to be explored.

E. 2017 CMS Display Measures

1. **Timely Receipt of Case Files for Appeals (Part D) & Timely Effectuation of Appeals (Part D).**

Summary of Comments: All commenters agreed to change the time frame of the Timely Receipt of Case Files for Appeals (Part D) & Timely Effectuation of Appeals (Part D) from the first six months of the current year to entire twelve months of the previous year.

Response: CMS will move forward with making this change.

2. **Medication Reconciliation Post Discharge (Part C).**

Summary of Comments: One third of the commenters were supportive of this measure being included on the 2017 display page and 2018 Star Ratings. However, almost half of the commenters had concerns about the timing of the measure and requested it delayed by either leaving it on the display page for at least two years or delay it on the display page and Star Ratings for at least a year. Many commenters wanted more clarification on the measure, specifically on the roles to complete the measure (i.e., social workers, pharmacists), data collection, and clarification or changes to technical specifications such as eligible members included in the denominator. There were also a few comments questioning the measure's construction, assumptions and validity. Examples included difficulty in collecting accurate information that medications were reconciled post discharge for D-SNP population, physicians coding accurately and addressing where members are utilizing multiple providers. A few thought the measure did not differ much from readmissions or the existent MRP measure required for physicians. Two comments requested benchmarking and cut point proposals including case mix and SES adjustments.

Response: CMS is planning to proceed to include this measure as part of the 2017 display page and the 2018 Star Ratings. This measure has been collected by SNPs for a number of years. Detailed specifications are available in HEDIS 2016, volume 2. We will monitor the 2016 data submissions for any data issues and modify our plans if needed.

3. **Hospitalizations for Potentially Preventable Complications (Part C).**

Summary of Comments: Almost all commenters were not supportive of the timing of the measure. They recommended it be delayed from the display page and Star Ratings or have it remain on display page for an additional year or two. Some commenters wanted information such as specifically what the ambulatory sensitive conditions are for this measure or requested

narrowing the scope on ambulatory sensitive conditions. A few of the comments had validity concerns about the measure specifically about the risk adjustment. One commenter had concerns with comparing the ESRD population with a general Medicare Advantage population in this measure.

Response: CMS is planning to proceed to include this measure as part of the 2017 display page and the 2018 Star Ratings. Detailed specifications are available in HEDIS 2016, volume 2. We will monitor the 2016 data submissions for any data issues and modify our plans if needed.

4. **Statin Therapy for Patients with Cardiovascular Disease (Part C).**

Summary of Comments: Almost all comments were negative, citing concerns with the measure's validity. Some noted a general lack of consensus with the 2013 ACC/AHA blood cholesterol guidelines, with others identifying detailed methodological issues. Several commenters noted that the measure does not account for members for whom statins are contraindicated, not well tolerated, not recommended, or refused. Others stated the measure does not account for alternate therapies or a wider range of statin dosages. A few commenters recommended moving this to a Part D measure like the related Statin Use in Persons with Diabetes (SUPD) measure and similarly excluding hospice beneficiaries. While a few agreed with the age ranges, just as many disagreed. Some commenters requested more specifics from CMS for the diagnostic codes and for "high or moderate statin." Because of concerns with the measure's validity, along with plans having limited time to implement quality improvement, several asked CMS to omit from the display page in 2017. Many others asked for the measure to stay on the display page for at least two years and/or to hold off reporting measure in the 2018 Star Ratings. Very few commenters agreed with CMS on this measure.

Response: Comments have been shared with NCQA. We are aware that treatments for cardiovascular disease are evolving and we will continue to monitor best practices for standards of care. We will keep the measure on the display page for an additional year to gain more experience with new treatment guidelines and metric and then add to the 2019 Star Ratings.

5. **Asthma Measures (Part C).**

Summary of Comments: The majority of commenters argued against adding the two proposed asthma measures as 2017 Star Rating display measures or as Star Rating measures in the future. These two measures expand NCQA's current asthma measures to include older adults and are defined on the basis of medication utilization.

The reasons against included: 1) belief that these are inappropriate measures for the majority of the Medicare population who, being ages 65 and over, are more likely to receive a COPD or other pulmonary disease diagnosis rather than an asthma diagnosis and comments argue for focusing efforts on areas that impact greater numbers of members; 2) concern about the difficulty for physicians to distinguish asthma from COPD in the senior population; 3) a few recent studies have indicated that medication management for people with asthma has not been shown to correlate with improved health outcomes, lower costs or lower utilization; 4) concern with the measures being constructed on the basis of medication utilization rather than on diagnosis given that some of the same medications are used for both asthma and COPD thus drug claims do not provide an accurate picture of which members have persistent asthma; 5) difficulty distinguishing between

'seasonal' versus 'persistent' asthma; 6) focusing on patients remaining on medications 'during the treatment period through the end of measurement year' requires defining an appropriate period of treatment for a condition which exhibits seasonality; 7) measures are not in line with NIH recommendations for step-down controller therapy or management of patients who exhibit seasonal variation in asthma symptoms; and 8) concern that for a population of persons often with multiple health conditions (seniors and the disabled), there are considerations of possible adverse medical consequences due to medication interactions.

A number of commenters requested CMS delay the addition of these 2 asthma measures for a few years until: 1) the measures are fully specified; 2) there is some evidence that asthma medication management is shown to correlate with improved health outcomes; 3) physicians are more experienced using the ICD-10 coding system which is thought to better delineate asthma from COPD; and 4) NCQA has fully specified, vetted and published results of these measures.

The only commenters arguing for the inclusion of the proposed asthma measures were pharmaceutical organizations.

Lastly, a number of comments argued that those plans which serve large numbers of the under age 65 Medicare population, the dual eligibles, would be further burdened by the inclusion of these measures due to socioeconomic factors which cannot be controlled but impact the frequency and severity of asthma events.

Response: CMS appreciates the comments received on this section. CMS shared the comments with NCQA and will continue to monitor development of these measures. CMS is planning to include these on the 2017 and possibly 2018 display page and will consider these for inclusion in Star Ratings for future years.

6. **Statin Use in Persons with Diabetes (SUPD) (Part D).**

Summary of Comments: Comments similar to those received for the Part C Statin Therapy for Patients with Cardiovascular Disease measure were submitted. There was mixed support to add the SUPD measure to the 2017 display page and to the 2018 Star Ratings. While there was support for this measure in general, a majority of commenters cited concerns with the measure's validity, or asked for the measure to stay on the display page for at least two years to gain experience or until methodological issues could be resolved. Several commenters noted that the measure does not account for members for whom statins are contraindicated, not well tolerated, not recommended, or refused. Others stated that the specifications should account for or exclude beneficiaries taking PCSK-9 therapies. Several commenters were concerned that prescription claims, not diagnosis codes, are used to determine the presence of diabetes, questioned the age criteria, or recommended excluding ESRD patients.

Response: Comments regarding the measure's technical specifications have been shared with the PQA. We will consider keeping the measure on the display page for an additional year to gain more experience with new treatment guidelines and metric and then add to the 2019 Star Ratings.

Forecasting to 2018 and Beyond

F. New Measures:

1. **Care Coordination Measures (Part C).**

   Summary of Comments: Support for new care coordination measures was mostly positive, although several commenters recommended against measures that involve chart review because of increased administrative cost. A few commenters requested more detail and for CMS to put specific proposed measures out for public comment before they become display measures. A couple of commenters recommended that CMS take into account mental health issues and dual/LIS/disabilities and consider a population-based approach. A few mentioned that measures should consider multiple provider types or suggested that CMS use existing reporting requirements related to care coordination. Several encouraged CMS to look at care coordination measures that are linked to improved outcomes. Some stated that it is critical for CMS to validate encounter data to ensure they are complete and accurate before relying on them.

   Response: CMS appreciates the comments received on this section. We shared comments with contractors developing care coordination measures and will continue to provide updates to the industry as the work progresses.

2. **Depression Measures (Part C).**

   Summary of Comments: Many commenters expressed concerns about privacy laws, as well as readiness of electronic systems to transmit clinical data from behavioral health providers. Several suggested CMS should first focus on depression screening measures before including a depression outcome measure. A few requested that other depression or mental health screening tools be included. A couple expressed concern that the 6 months measurement window is too short to demonstrate impact. Several stated that the measure should be on display for several years following HEDIS approval.

   Response: CMS shared comments received on this topic with NCQA and will continue to monitor development of the measure. NCQA is also working on a Depression Screening and Follow-Up measure which may be proposed for HEDIS in the future.

3. **Appropriate Pain Management (Part C).**

   Summary of Comments: NCQA is exploring opportunities to develop a new measure focusing on appropriate pain management. The intent is to assess the quality of pain management and treatment.

   Commenters expressed appreciation for the exploration of this important topic. However, commenters stressed that the experience of pain is subjective and varies across individuals, conditions, and stages of a condition. Commenters stressed that these considerations need to be addressed in specifying an appropriate pain management measure so that valid comparisons can be made across plans and member populations.

Commenters requested full measure specification to enable plans to provide meaningful comment. Further, that if a measure is created and implemented, it remain on the display page for at least 2 years.

One commenter suggested that a standardized pain screening tool be developed / employed to assess, document, and monitor the experience of pain. Another commenter suggested that in terms of pain control, the measure address not only medication use but also alternative treatments for managing pain. Lastly, two commenters indicated that appropriate pain management can be at odds with controlling/monitoring for opiate use.

Response: CMS appreciates the comments received on this section. CMS shared the comments with NCQA and will continue to monitor development of these measures.

4. **Use of Opioids from Multiple Providers or at High Dosage in Persons without Cancer (Part D).**

Summary of Comments: There was general agreement on the importance of resolving opioid overutilization and not adding the measures to the Star Ratings at this time. While there was some support to add these three PQA opioid overutilization measures to the 2018 Part D display page (using 2016 data), most commenters were opposed or expressed concerns with the proposal. Concerns about the proposal included: 1) neither clinical guidelines nor experience support the validity of the measures, 2) the measures do not measure and will not support plan performance, particularly #1 and #2, 3) more detailed descriptions of the measure specifications are needed in order to consider support, 4) some sponsors will be disadvantaged based upon the characteristics of their enrollees (e.g., high numbers of disabled enrollees), 5) Part D sponsors have limited ability to influence prescriber behavior, and 6) if the measures are not suitable for Star Ratings, they should not be display measures.

Response: Due to concerns raised by commenters, CMS will implement the three PQA opioid overutilization measures as Patient Safety measures, not display measures, for one year to gain experience with the measures and pending new guidelines (e.g., from CDC) and current research on opioid prescribing, overutilization, and interventions. We shared specification comments with the PQA.

5. **Antipsychotic Use in Persons with Dementia (APD) (Part D).**

Summary of Comments: There was general agreement that this measure addresses an important issue. Over half of the commenters agreed with the proposal, and one-third were neutral. Changes to the proposal were suggested by two commenters, and over two-thirds of the commenters noted specific concerns, including 1) lack of access to diagnosis data required for the measure, 2) limitations on the ability of sponsors to intervene because antipsychotic drugs have protected class status, 3) a desire to review the complete measure specifications, and 4) this measure is primarily associated with nursing homes, the nursing home quality rating reporting system includes a related measure, and the facilities are responsible, not plans.

Response: CMS will proceed as proposed. CMS will include a link to the CMS.gov webpage for an APD measure analysis report that provides detailed specifications and testing results. We shared specification comments with the PQA.

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

1. **Colorectal Cancer Screening (Part C).**

   Summary of Comments: This measure is under consideration for revision. The USPSTF will release its revised guidelines in late 2016. At that time, NCQA will consider revising the Star Rating Colorectal Cancer Screening HEDIS measure.

   The comments supported waiting for the USPSTF's final evidence-based recommendations for colorectal cancer screening methods. The commenters stressed the need for plans to be provided, in advance of implementing any changes to the Star Ratings, additional detail of specification changes to allow stakeholders to provide meaningful comment. The comments also stressed the need for advance notice of the timeline for implementing changes so as to educate providers and beneficiaries. It was suggested that any revised measure not be implemented until the 2019 Star Rating year.

   One commenter requested there be no change to the measure's age limits. One commenter requested the measure be broken down by age groups. Another commenter stressed that universal colorectal cancer screening is not supported for persons with ESRD.

   Response: CMS appreciates the comments received on this section. CMS shared the comments with NCQA and will continue to monitor development of these measures.

2. **Fall Risk Management (Part C).**

   Summary of Comments: Commenters were equally split between supportive, negative and neutral comments. Many commenters would like more time for research or for plans to prepare for changes. Some commenters would prefer CMS not use measures derived from surveys of beneficiaries. A large number of commenters suggested that while it was appropriate to update specifications to reflect changing guidelines, there is still not enough evidence to consider vitamin D as a treatment to reduce falls. Others commented that specification changes can change the focus of the measure and "eliminate its principal focus of prevention of falls among members by eliminating physical barriers." A few Special Needs Plans serving dialysis patients expressed concerns that vitamin D is not appropriate treatment for patients also getting dialysis (it risks leading to hypercalcemia) and therefore some exclusion is needed for those plans or for most of their enrollees.

   Response: CMS shared these comments to NCQA and highlighted the concerns commenters have about the evidence for use of vitamin D as a treatment, as well as the concern that Special Needs Plans that focus on End Stage Renal Disease may not have enrollees appropriate for this measure.

3. **Pneumococcal Vaccination Status for Older Adults (Part C).**

   Summary of Comments: Almost all comments were negative, citing concerns that the measure relies on Medicare CAHPS survey data. Self-reports, according to commenters, are unreliable because members may not recall getting a specific vaccine, especially if administered a long time ago. Instead, most commenters recommended CMS use claims data instead. Still, a few cautioned that claims data is imperfect and ask for lower cut points. Some plans asked CMS to

keep the measure as display and not add to Star Ratings until tested further, while a pharmaceutical company encouraged CMS to include it the Star Ratings.

Response: We appreciate the concerns of commenters. We will monitor data during the display period for issues with validity.

4. **CAHPS measures (Part C & D).**

Summary of Comments: Commenters were mostly positive about the proposed change to the 5.0 version of CAHPS. Several requested that the CAHPS survey be shortened in order to increase response rates and stated that questions should focus on questions related to consumer satisfaction. A few commenters expressed concern with CAHPS in general as they feel it is subjective. One expressed concern about benchmark rate changes and CAHPS scores. Comments on the sampling proposal were mixed.

Response: Patient experience surveys such as CAHPS focus on how patients experienced or perceived key aspects of their care, not how satisfied they were with their care. CAHPS surveys follow scientific principles in survey design and development. The surveys are designed to reliably assess the experiences of a large sample of patients. They use standardized questions and data collection protocols to ensure that information can be compared across healthcare settings. CAHPS surveys are developed with broad stakeholder input, including a public solicitation of measures and a technical expert panel, and the opportunity for anyone to comment on the survey through multiple public comments period through the Federal Register. Regarding survey length, analyses suggest that the relationship between survey length and response rate for the MA CAHPS survey is only weakly negative. Specifically, the use of 12 supplemental items as compared to none was associated with a 2.5% reduction in response rate (Beckett et al, Public Opinion Quarterly, in press)[1].

To examine the impact of decreasing benchmarks on CAHPS measure scores, we conducted an analysis of changes in CAHPS stars with changes in proxy standardized MA benchmarks at the contract level. The benchmarks were developed as an average of plan-level BPT benchmarks weighted by July enrollment for each year. Analyses on the impact on CAHPS scores did not find a systematic relationship between CAHPS scores and declining benchmarks.

The table below includes a comparison of contracts changing their CAHPS measure star rating from 2014 to 2015 comparing the quartile with the biggest drop in benchmarks ($81 or more) with the quartile with the smallest changes in benchmarks (+$30 to -$47):

---

[1] Beckett MB, Elliott MN, Gaillot S, Haas A, Dembosky JW, Giordano LA, Brown J. (In Press) "Establishing limits for supplemental items on a standardized national survey." Public Opinion Quarterly

Measure	Quartile with largest drop in benchmarks (% increasing stars/% staying same/% decreasing)	Quartile with smallest change in benchmarks (% increasing stars/% staying the same/% decreasing)
Care Coordination	23%/51%/26%	24%/51%/25%
Customer Service	22%/45%/33%	30%/45%/25%
Getting Appointments and Care Quickly	21%/61%/18%	11%/68%/21%
Getting Needed Care	23%/46%/31%	23%/48%/29%
Rating of Health Care Quality	22%/54%/24%	9%/69%/21%
Rating of Health Plan	18%/60%/22%	20%/56%/24%
Getting Needed Prescription Drugs	18%/45%/36%	20%/51%/29%
Rating of Drug Plan	21%/57%/21%	24%/59%/17%

5. **Medication Adherence for Hypertension (RAS Antagonists) (Part D).**

Summary of Comments: All commenters supported the proposed exclusion. One commenter requested we exclude dual-eligible beneficiaries that have $0 copay. Another requested we emphasize ESRD patients are excluded. A few requested CMS clarify which year this change applies; two requested CMS set upper age limits for all 3 medication adherence measures.

Response: We will proceed as planned. We will clarify that this exclusion will be applied for the 2017 Star Ratings. We shared specification change suggestions with the PQA.

6. **MPF Price Accuracy (Part D).**

Summary of Comments: About one-quarter of commenters agreed with the proposed changes, and the remaining were neutral or opposed. The most common comments included:

- Broaden the cut points because the measure scores are tightly clustered resulting in insignificant/negligible differences and therefore do not inform beneficiaries of any impactful differences among the plans.

- Retire or move the measure to the display page due to high scores.

- Change the methodology for scoring this measure.

- Clarify CMS' decision making process that led to this change, and provide details of our analysis of this change.

- Provide best practices for how to achieve a high score.

- Move to display page for one year due to methodology changes.

Other commenters commented that the different frequency of point of sale and MPF pricing updates is a barrier to improving plan performance. Also, if the measure has new scoring methodology, commenters suggested temporarily removing the measure from the Improvement

Measure. Lastly, some commenters would like to be able to access detail reports of their scores earlier and more frequently.

Technical comments included 1) the amount of difference between PDE prices and MPF prices that constitute an inaccuracy should be broadened to more than one half of one cent, 2) only the Formulary Reference File NDCs should be used when selecting the PDE claims that are being measured because pricing for the same drug, strength, and dosage form can vary from manufacturer to manufacturer, and 3) the "Patient Residence Code" should be used to determine the retail claims from the PDEs.

Response: We will proceed with these changes as planned. CMS' simulations found that the accuracy scores using the new methodology were generally similar to scores calculated using the current methodology. This measure will continue to be excluded from the Improvement Measure. CMS uses accepted mathematical algorithms and practices to formulate the thresholds (cut points) required to earn a particular rating. Currently the majority of contracts receive high Accuracy scores. We do not believe options to further differentiate plans' performances would be supported by sponsors, as this could entail identifying price differences smaller than one half of one cent. We continue to be open to other changes that will improve the impact of this measure.

7. **Drug-Drug Interactions (DDI) (Part D Display).**

Summary of Comments: Most commenters were neutral, while a few voiced concern about the implementation of future changes to PQA specifications including the timing of adoption by CMS for the display measure.

Response: We will proceed as planned. Any future changes will be announced via the annual Request for Comments and draft Call Letter process to provide advance notice.

8. **Center for Medicare and Medicaid Innovation Model Tests.**

Summary of Comments: Commenters appreciated CMS' attention to the potential for improvements in quality in the MA-VBID and Part D Enhanced MTM models to favorably influence Star Ratings for contracts with participating plans and expressed a desire for contracts not to be disadvantaged for either participating or not participating. Several requested that CMS provide additional information about impact on plans. Several mentioned that only some contracts are permitted to participate in the models, a few suggested that SNPs and territories (e.g., Puerto Rico) should be allowed to participate. Some suggested that waiving MTM reporting requirements for participating plans could impact cut points for this measure, and a few requested that CMS suspend the MTM Star Rating during the model test years.

Response: CMS will closely monitor performance of contracts participating in the model to evaluate any effect on Star Ratings. Our goal is to ensure that contracts are not penalized. Some possible options are to establish different cut points for model participants and to case mix adjust scores for the purpose of determining cut points. We will provide more information to stakeholders as it is available and continue to engage with stakeholders on the impact of these models.

H. *Measurement and Methodological Enhancements.*

Summary of Comments: Comments ranged in topics from general to measure-specific. They included comments about specific Star Rating changes and measures, cut points, development of outcome measures, the Reward Factor, as well as display measures.

Response: CMS appreciates all comments and will explore the feasibility of specific proposals for possible future implementation. For example, we will continue to look at the issue of whether to conduct call center monitoring in languages proportional to the prevalence of each language in the 65 and older U.S. population.