RAND Medicare Advantage (MA) and Part D Contract Star Ratings Technical Expert Panel May 31st 2018 Meeting

PRESENTATION

Cheryl L. Damberg and Susan M. Paddock
Overview of the Part C and Part D Star Ratings

RAND Technical Expert Panel
Star Ratings for Medicare Advantage and Prescription Drug Plans

Susan Paddock, PhD
May 31, 2018
Topics Addressed

• Purpose of Star Ratings
• Overview of the Star Ratings methodology
• Enhancements to the methodology
PURPOSE OF STAR RATINGS
Purpose of Star Ratings

• Improve the level of accountability for the care provided by physicians, hospitals, health plans, and other providers

• Stimulate improvement in care through public accountability and performance-based incentives (financial, marketing/enrollment)

• Assist beneficiaries in finding the best plan
Ratings on Medicare Plan Finder (MPF)

Plan Quality and Performance Ratings

Choose Plans to Compare
When you choose 3 plans to compare, quality and performance information will be available to help you make the best choice for you. Quality and Performance varies across plans. Giving good quality care means doing the right thing, at the right time and in the right way to get the best results possible.
Medicare Plan Finder: Your Results Page

Symbols

Star: When you see this symbol near a plan name, it means that Medicare Program gave the plan a 5-star (the highest rating). If a plan has a 5-star rating, people with Medicare can switch into that plan at any time during the year, even if it’s not during an enrollment period.

Red Exclamation Point: Where you see this icon next to a plan, it means that Medicare has given the plan a low health or drug plan summary rating (or both) for 3 years in a row. If you are considering enrolling in such a plan, look closely at the plan’s ratings for specific topics.
High Performing Plans

- CMS highlights contracts receiving an overall rating of 5 stars

- Beneficiaries may enroll in a 5-Star PDP, MA-PD, or MA-Only plan through a Special Election Period (SEP)

- 5-Star contracts allowed to market year-round
Consistently Low Performing Plans

- Icon displayed for contracts rated less than 3 stars for at least the last 3 years for their Part C or D rating.
- Beneficiaries may not enroll online via the MPF in a low-performing icon (LPI) plan. Beneficiaries must contact the plans directly.
- Notices are sent to beneficiaries in LPI plans explaining they are in a low performing plan and eligible to move to a higher quality plan.
MA Contract Quality Bonus Payments

• Per the Affordable Care Act, CMS makes Quality Bonus Payments (QBPs) to MA contracts that meet quality standards as indicated by the Star Ratings

• MA contracts with overall Star Ratings of 4 or greater are eligible for QBPs
OVERVIEW OF STAR RATINGS METHODOLOGY
Measure Development

- CMS looks to consensus building entities such as National Committee for Quality Assurance (NCQA) and Pharmacy Quality Alliance (PQA) for measure concept development and specifications
  - Most measures have achieved NQF endorsement

- Measure set reviewed each year; move towards more outcome measures

- Measures that move from the Star Ratings to CMS’ display page still used for compliance and monitoring
Star Ratings Cover 9 Domains
(46 unique measures across Parts C & D)

<table>
<thead>
<tr>
<th>Ratings of Health Plans (Part C)</th>
<th>Ratings of Drug Plans (Part D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staying healthy: screenings, tests, vaccines</td>
<td>Drug safety and accuracy of drug pricing</td>
</tr>
<tr>
<td>Managing chronic (long-term) conditions</td>
<td>Member complaints and changes in the drug plan’s performance</td>
</tr>
<tr>
<td>Member experience with the health plan</td>
<td>Member experience with the drug plan</td>
</tr>
<tr>
<td>Member complaints and changes in the health plan’s performance</td>
<td>Drug plan customer service</td>
</tr>
<tr>
<td>Health plan customer service</td>
<td></td>
</tr>
</tbody>
</table>
Measure Weights

• The Star Ratings measures span five broad categories:
  – Improvement – 5
  – Outcomes/Intermediate Outcomes – 3
  – Patient experience – 1.5
  – Access – 1.5
  – Process – 1

  *Each measure is assigned a weight using category definitions included in the Star Ratings Technical Notes*

• Patient experience and access measure weights will increase to 2.0 with the 2021 Star Ratings
Methodology for Star Assignments: Two Techniques

• Clustering

• Relative Distribution and Significance Testing
Clustering

- Assigns cut points by minimizing differences within star categories and maximizing differences across 5 star categories.
Relative Distribution and Significance Testing

- Consumer Assessment of Healthcare Providers and Systems (CAHPS) Survey
- Takes into account the relative percentile distribution and reliability with significance testing (whether it is statistically significantly higher or lower than the national average)
Structure of the Star Ratings

Overall
Weighted Average
(whole & ½ stars)

Summary
Weighted Average
(whole & ½ stars)

Domain
Non-weighted Average
(whole stars)

Measure
(numeric values & whole stars)

Overall Rating
MA-PD

Part C Rating
MA-Only & MA-PD

Part D Rating
MA-PD & PDP

HD1 HD2 HD3 HD4 HD5
DD1 DD2 DD3 DD4

C01 . . . . . . . . . . . . . . . . . . . Cnn

D01 . . . . . . . . . . . . . . . . . . . Dnn
Final Overall and Summary Star Ratings include two additional factors which are added

1) Reward factor
   - Calculate the mean and variance of the individual performance measure stars at the contract level
   - Contracts with means greater than or equal to the 65th percentile and variances less than the 70th percentile receive reward factors of 0.1-0.4 stars

2) Categorical Adjustment Index (CAI)
   - Accounts for within-contract differences in performance, based on percentages of low-income and disabled beneficiaries
   - Assigned based on the contract’s % disabled and % receiving a low-income subsidy and/or dually eligible
   - An interim solution while measure developers evaluate whether and how to adjust performance measures for patient sociodemographic characteristics
ENHANCEMENTS TO THE METHODOLOGY
Goals for Star Ratings Enhancements

CMS continuously reviews the Star Ratings methodology and seeks to enhance it to:

• Improve the process and transparency surrounding the calculations
• Incentivize plans to foster continuous quality improvement in the MA and Part D programs
• Provide information that is a true reflection of the quality of care provided
• Provide stable ratings over time
• Treat contracts fairly and equally
• Minimize unintended consequences
Recent Enhancements

- 2014 Star Ratings (Fall 2013)
  - Increased the weight of improvement measures from 1 to 3

- 2015 Star Ratings (Fall 2014)
  - Increased the weight of improvement measures from 3 to 5, to further reward contracts for the strides they have made to improve care, especially for those serving challenging populations
Recent Enhancements (continued)

- 2016 Star Ratings (Fall 2015)
  - Eliminated pre-set 4-star measure thresholds
  - Added contracts with enrollment from 500-999 to Star Ratings

- 2017 Star Ratings (Fall 2016)
  - Added Categorical Adjustment Index: interim analytical adjustment for socioeconomic/disability status
Methodology enhancements and multi-stakeholder input:

Call letter process

- Transparency and multi-stakeholder input

- Annual Call Letter process includes a request for comments on proposed Star Ratings enhancements
  - Draft Call Letter: early February
  - Final Call Letter: early April

- Comments received from multiple stakeholders
  - professional organizations
  - Medicare Advantage (MA) and Part D sponsors
  - advocacy groups
  - state Medicaid agencies
  - pharmaceutical manufacturers
  - pharmacy benefit managers
  - pharmacies
  - concerned citizens
Methodology enhancements and multi-stakeholder input:

Regulatory process

- Notice of Proposed Rulemaking – November 2017
  - Multiple stakeholders submitted comments during the comment period
- Final Rule, Federal Register – April 2018
- Rule will be updated when new measures, measures with substantive changes, or methodology changes are proposed
Considerations for How to Compute the Measure Cut Points Used to Assign Measure Stars

RAND Technical Expert Panel
Star Ratings for Medicare Advantage and Prescription Drug Plans

Susan Paddock, Marc Elliott, Megan Mathews, Feifei Ye, Cheryl Damberg

May 31, 2018
Outline

• Background
• Review current methodology for determining cut points used to assign measure stars
• Summarize desirable attributes of measure cut points
• Measure score trend analysis
• Questions for the Technical Expert Panel’s consideration
Background

- CMS is revisiting the methodology used for setting cut points and assigning measure stars for the majority of Star Ratings measures

- Interest focuses on modifications that provide:
  - Greater stability in cut points year over year
  - Advance notice of cut points
Goals of CMS Star Ratings Enhancements

• Improve the process and transparency surrounding the calculations

• Incentivize plans to foster continuous quality improvement in the MA and Part D programs

• Provide information that is a true reflection of the quality of care provided and the experiences of beneficiaries in the plan

• Provide stable ratings over time

• Treat contracts fairly and equally

• Minimize unintended consequences
Current Methodology for Star Assignments

• Clustering
  – HEDIS, HOS, prescription drug event (PDE) data, disenrollment rates, and CMS administrative data-based performance measures

• Relative Distribution and Significance Testing
  – Consumer Assessment of Healthcare Providers and Systems survey-based performance measures
  – *Consumer Assessment of Healthcare Providers and Systems measures are not the focus of today’s discussion*
**Goal:** Separate the contract measure scores into 5 non-overlapping groups (clusters) to determine the measure cut points.
Clustering

• Approach: Minimize distances of scores within each cluster and maximize distances across clusters
  – Hierarchical (agglomerative) clustering

• *Cut points (or thresholds)*: the set of values that break the distribution of scores into 5 non-overlapping clusters
  – Determined as the measure score minimums of the four best performing clusters (for ‘higher is better’ measures)

• Each cluster will be assigned a *measure star*
  – Five measure star values: 1, 2, 3, 4, 5
Medication Adherence Cut Points, 2015-2018
Examples of Factors that Can Influence Annual Changes in Cut Points

- Newer contracts with lower performance
- Structural changes that drive substantial gains year-to-year
- New care processes that allow for more rapid improvement
- Poorer capture of the data needed to demonstrate performance
Examples of Cut Point Trends Over Time

Desirable Attributes of Cut Points

- Accurate and meaningful reflection of performance trends
- Stable over time
- Provide advance notice of changes in cut points
Concerns about Stability of Cut Points

- Clustering can be sensitive to outliers
Concerns about Stability of Cut Points

- Improving/maintaining physical health, histograms
Concerns about Stability of Cut Points

- Clustering can be sensitive to outliers
Concerns about Stability of Cut Points

- Adult BMI Assessment

![Graph for 2017 and 2018 with contract score=11]
Concerns about Stability of Cut Points

A contract could improve performance but lose a measure star (illustrated by dashed lines)
Cut Point Concern: Advance Notice

- Performance targets not known until after the measurement period
- Example: For the 2019 Star Ratings, the clustering approach will be applied to measure data collected during 2017
How Might Advance Notice Be Provided?

• Use prior Star Ratings data
  – How should the data be used?
    • Should cut points be predicted using prior data?
    • How might limits on cut point change be determined?
  – Which years of data should be used?

• Set absolute cut points using other criteria
  – How should these be determined?
# Timeline for Hypothetical Types of Advance Notice: 2021 Ratings Year

<table>
<thead>
<tr>
<th>Measurement Year (MY) for Star Ratings Year (RY) 2021</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO ADVANCED NOTICE: Current approach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>PARTIAL ADVANCED NOTICE: Lag cut points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARTIAL ADVANCED NOTICE: Project cut points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FULL ADVANCED NOTICE: Lag cut points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FULL ADVANCED NOTICE: Project cut points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **2016**: Performance data for RY 2021 cut points
- **2017**: Performance data for RY 2021 cut points
- **2018**: Performance data for RY 2021 cut points
- **2019**: Performance data for RY 2021 cut points
- **2020**: Cut Points Calculated: Announced in September
- **2021**: Cut Points Calculated: Announced in September

- **Ratings Year**: Cut points are known benchmarks for MY 2019 data collection

---

*Slide 43*
Measure Score Trends, 2014-2018

• General findings
  – Improvements in performance for most measures
  – Pattern of improvement varies by measure
  – Decreasing variation in many measure scores over time
Trends For Different Types Of Measures

• Electronic health record-enabled

• Relatively new measures with greater opportunity for improvement (i.e., lower performance, more variation)

• Outcome measures
Electronic Health Record-enabled Measures

- CMS provided incentives for eligible providers to adopt certified electronic health records and report quality measures under the Meaningful Use initiative.

- Moving between Meaningful Use 1 (2014) and Meaningful Use 2 (2016), CMS proposed a recommended core set of measures which included measures assessed in Star Ratings program (e.g., BMI assessment).

- Adoption of electronic health records functionalities can lead to substantial improvements year-to-year.
Electronic Health Record-enabled Measures

KEY

75th percentile

25th percentile

Boxes represent the interquartile range (IQR)

Adult BMI Assessment

max
median
mean
min

year

2014  2015  2016  2017  2018
Electronic Health Record-enabled Measures

Adult BMI Assessment

Scores improve during 2014-2016, the same time period as Meaningful Use 1-Meaningful Use 2
Electronic Health Record-enabled Measures

Adult BMI Assessment

Inter-quartile range of scores is narrowing over time
Electronic Health Record-enabled Measures

COA Functional Status Assessment (for SNPs)

COA Pain Assessment (for SNPs)

Improvement during Meaningful Use 1 - Meaningful Use 2
Electronic Health Record-enabled Measures

*Greater improvements toward lower end of measure score distribution – and decreasing IQRs*
Relatively New Star Ratings Measures

• Performance on new measures typically leaves more room to improve

• Large year-to-year gains are possible and desirable from a quality improvement perspective
Outcome Measures

- Typically little change and similar variation over time
Stable Measure Distributions

- Some measure score distributions are stable over time, changing little or with small and steady improvements
  - Screening (breast cancer, colorectal cancer)
  - HOS maintaining/improving physical health, monitoring physical activity
  - HEDIS diabetes care measures
  - Rheumatoid arthritis treatment
  - Reducing the risk of falling
  - Call center / foreign language / TTY availability
  - Appeals upheld
  - Members choosing to leave the plan
  - Medication adherence measures
Measure Score Trends: Summary

- Mean performance is improving over time for most measures
  - Improvements are greater toward the lower end (e.g., 25\textsuperscript{th} percentile) versus higher end (e.g., 75\% percentile) of the distribution

- There is decreased measure score variation over time

- Structural changes can lead to changes in a short time period

- Older data do not often accurately reflect current performance

- Trends differ across measures
Questions for the TEP’s Consideration

- What types of approaches should be considered to reduce significant year-over-year changes in thresholds used to assign stars?

- How should CMS balance using the most current data to set cut points versus applying several years of older data to allow plans to have advance notice of thresholds prior to the start of a measurement year?

- Should absolute thresholds be considered for some measures?
  - If yes, how should they be determined and change over time?
Questions for the Technical Expert Panel’s Consideration

• Which approaches would be most promising to yield these desirable attributes of cut points?
  – Accurate and meaningful reflection of performance trends
  – Stability over time
  – Advance notice of cut points
Utility and Feasibility of Constructing and Reporting Star Ratings at the Level of Plan Benefit Package or Geographic Area

Marc N. Elliott, Justin Timbie, Cheryl Damberg

May 31, 2018
Potential Levels for Measurement and Reporting

- **Contract level**
  - Current approach used for reporting Star Ratings and QBPs to MA contracts

- **Smaller, standardized geographical units within contracts**
  - Contracts may have service area that covers multiple states or regions

- **Plan benefit package (PBP) level**
  - Each contract may have 1 or many PBPs from which beneficiary may choose
Simplified Example of Different Types of Contracts

(mean # PBPs/contract = 7.25, median = 4.0, 25\textsuperscript{th} - 75\textsuperscript{th} = 1.0 - 8.0 PBPs)

DSNP, which refers to a Dual Eligible Special Needs Plan, is a type of PBP
REPORTING USING SMALLER, STANDARDIZED, GEOGRAPHIC UNITS WITHIN CONTRACTS
Contracts Vary in the Size and Spread of their Service Areas

Based on 2016 data
Small Sample Sizes Would Be a Challenge for Geographic Reporting

Contract H3931:
- HMO
- 12 States in Service Area
- Total enrollment: 53,911

Based on 2016 data
# Commonly Used Geographic Units in Health Services Research

<table>
<thead>
<tr>
<th>Geographic Unit</th>
<th>Description</th>
<th>Aggregations of</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital Referral Region</td>
<td>HRR</td>
<td>Tertiary care markets</td>
<td>Zip Codes</td>
</tr>
<tr>
<td>Metropolitan Statistical Area</td>
<td>MSA</td>
<td>Areas of high population density</td>
<td>Counties</td>
</tr>
<tr>
<td>Health Service Area</td>
<td>HSA</td>
<td>Hospital care markets</td>
<td>Counties</td>
</tr>
<tr>
<td>Primary Care Service Area</td>
<td>PCSA</td>
<td>Primary care markets</td>
<td>Zip Codes</td>
</tr>
</tbody>
</table>
MedPAC Recommended Smaller, Standardized Geographic Units for Quality Reporting

• Quality reporting should use geographic units that are "meaningful for beneficiary decision making" and "accurately reflect local patterns of care delivery"

• MedPAC has recommended using MSAs and HSAs, but also suggested exploring other geographic units

*MSA: Metropolitan Statistical Area
*HSA: Health Service Area*
MedPAC recommends calculating star ratings for each contract in each unit for public reporting and QBPs.
PLAN BENEFIT PACKAGE LEVEL REPORTING
CMS Recently Solicited Comments on PBP-Level Reporting

• Ability to distinguish heterogeneous plan benefit offerings
  – Example: Beneficiary enrolled in a DSNP in a mixed contract (i.e., dual and non-dual PBPs within same contract)
    • Contract-reporting: performance is enrollment-weighted across PBPs
    • PBP-reporting: DSNP performance in mixed contract compared to DSNP performance offered by a different contract

• PBP-level reporting may preclude further geographic disaggregation
  – Contract-reporting: might be able to drill-down to mixed contract’s performance in LA County
  – PBP-reporting: DSNP performance may be for all of California

• Commenters raised concerns about burden, sample size issues
One Approach: PBP Replaces Contract as Reporting Unit

- Likely requires substantial increased data collection to ensure enough PBPs are represented
  - Larger Consumer Assessment of Healthcare Providers and Systems (CAHPS) samples with somewhat different sample design
  - Census sampling for 6/25 HEDIS measures that currently use the hybrid approach in 2018
  - All/smallest PBPs probably never possible

- Micro-geographic-level reporting may not be possible for separate PBPs
  - But enrollment weighting can fix gaming associated with mergers/consolidations; CMS to implement starting with 2020 Star Ratings and QBPs
A Second Approach: PBP Reporting

Supplements Contract-Level Reporting

- Could be a drill-down option to inform choice and monitor quality, but with no change to Star Rating
  - Where sample size is sufficient (also for geographic-unit reporting)
- Would happen independently of any other changes that may be made to geography
- Does not require (but could be done better with) additional data collection
- Concern that PBP reporting not possible for many PBPs
  - Those in smallest PBPs and contracts surveyed in successive years
**Metrics for Evaluating Alternative Reporting Units**

**Intra-class correlation coefficients** (ICCs) are the proportion of variance in patient-level scores associated with the reporting unit. Determines how big of a sample size you need for good reliability.

**Reliability** is the proportion of variance in PBP- or contract-level scores measures associated with the reporting unit for a given sample size; measures how well data distinguishes performance of the reportable units. Ideally at least 0.8 for p4p and 0.7 for reporting.

**Informativeness** is the percent of variance of PBP scores not predictable from contract scores; when at least moderate informativeness (0.33) indicates that PBP-reporting provides meaningful information beyond what is known from contract-reporting.
Summary of Results of Analyses of PBP-Level Reporting

<table>
<thead>
<tr>
<th></th>
<th>CAHPS</th>
<th>HEDIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBP-level ICCs</td>
<td>As high for PBPs as for contracts</td>
<td>As high for PBPs as for contracts</td>
</tr>
<tr>
<td>Reliability &gt;0.80 at 100 survey completes</td>
<td>All measures</td>
<td>All measures</td>
</tr>
<tr>
<td>20+% of PBPs have 100+ completes (~250 enrollees CAHPS)</td>
<td>8 of 9 measures</td>
<td>10 of 14 measures</td>
</tr>
<tr>
<td>Majority of Contracts Have at Least 1 PBP with 100+ Completes</td>
<td>7 of 9 measures</td>
<td>11 of 13 measures</td>
</tr>
<tr>
<td>20%+ Contracts Have at Least 2 PBPs (of an average of 7) with 100+ Completes</td>
<td>7 of 9 measures</td>
<td>10 of 13 measures</td>
</tr>
<tr>
<td>Moderate to Strong Informativeness*</td>
<td>All measures</td>
<td>4 of 13 measures</td>
</tr>
</tbody>
</table>

*Percent of variance of PBP scores not predictable from contract scores.
Questions for the TEP (1 of 2)

- What analyses should RAND consider performing to inform an understanding of the feasibility and value of reporting Star Ratings at “other units of measurement”?

- If contract reporting occurred at a standardized geographic level, what geographic unit should be considered?

- If reporting occurred at the level of the PBP, what types of comparisons should be considered?
  - Comparisons of similar PBPs across contracts (like two different D-SNP PBPs embedded within contracts with non-SNP parts) or
  - Comparisons of different PBPs within a single contract?
Questions for the TEP (2 of 2)

- Are there approaches to identifying and addressing noncontiguous contract service areas that do not require geographic drilldowns of all contracts?

- Is it acceptable to increase sample sizes and costs for Consumer Assessment of Healthcare Provider Systems, HOS, and HEDIS hybrid method sampling to allow PBP and/or geographic stratification? Should census sampling be required to increase HEDIS sample sizes?
  - To supplement contract-level, or replace contract-level? (large costs and less gain for HEDIS)

- How to handle PBPs or geographic reporting units with not enough enrollees to reliably measure performance?
Measures Used to Assess Medicare Advantage and Prescription Drug Plan Contract Performance

RAND Technical Expert Panel
Star Ratings for Medicare Advantage and Prescription Drug Plans

Rachel Reid, MD, MS & Cheryl Damberg, PhD

May 31, 2018
Outline of Topics for Discussion

• Review existing measures

• Identification of new measures
  – Topic areas, measure types, data sources
  – Potential analyses for measures using existing administrative data sources

• Considerations for analyses regarding potential removal of measures
Existing Measures

• 2019 Star Ratings Measures by Domain, Type, and Data Source

• 2019 Display-Only Measures, which include several categories:
  – No longer contributed to 5-Star Rating calculation, still monitored/measured for informational purposes only
  – Removed from 5-Star Ratings Calculations temporarily due to a substantive specifications change
  – New measures that have not previously contributed to 5-Star Ratings calculation

• 2 New Star-Ratings Measures for 2019, Previously Display-Only Measures
  – Statin Use in Persons with Diabetes
  – Statin Therapy for Patients with Cardiovascular Disease
As RAND plans analyses related to identification of potential candidate measures, should we consider:

<table>
<thead>
<tr>
<th>Particular topic areas or domains?</th>
<th>General measure types?</th>
<th>Particular measures?</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g., Care Coordination, Vulnerable/High-need/Multi-Morbid Patients, Low-Value Care or Efficient Use of Resources</td>
<td>e.g., Process, Intermediate Outcome, Outcomes, Patient Reported Outcomes, Access, Patient Experience, Plan Operations</td>
<td></td>
</tr>
</tbody>
</table>
Considering Use of Existing Data Sources

Data collection and reporting for quality measurement has associated burden

- Opportunities may exist to marshal existing Administrative Data sources (e.g., Encounter Data)

Questions:

- Are there specific measures or topic areas using Administrative Data that RAND should consider?
- What considerations or criteria should RAND use when identifying or analyzing measures using Administrative Data?

- RAND conducted a scan to identify measures calculable via Encounter Data with other administrative data (e.g., PDE data) and applicable to Medicare Advantage and Part D Plans:
  - Identified measures across several areas: Appropriate Medication Use, Medication Monitoring and Adverse Effects, Ambulatory Care Sensitive Conditions, Resource Use, Chronic Disease Management, Care Coordination and Follow-up

- Examples:
  - Use of opioids at high dosage and use of opioids from multiple providers
  - Readmissions from Post-Acute Care
(e) Removing measures.

1. CMS will remove a measure from the Star Ratings program as follows:
   a. When the clinical guidelines associated with the specifications of the measure change such that the specifications are no longer believed to align with positive health outcomes;
   or
   b. A measure shows low statistical reliability.

2. CMS will announce in advance of the measurement period the removal of a measure based upon its application of this paragraph (e) through the process described for changes in and adoption of payment and risk adjustment policies in section 1853(b) of the Act in advance of the measurement period.
Removing Measures from Star Ratings

What criteria might RAND use in analyses surrounding transition of existing Star Rating Measures to Display-Only Measures?

<table>
<thead>
<tr>
<th>Lack of meaningful variation?</th>
<th>Attainment of a threshold average score across plans?</th>
<th>Lack of improvement over time?</th>
<th>Other Criteria?</th>
</tr>
</thead>
</table>

Examples of current Display-Only Measures previously removed/retired from the Star Ratings are denoted in the hand-out by the year in which they were moved to Display-Only status.
For Analyses Related to Potentially “Topped Out” Measures

What options should RAND consider in analyses regarding defining a measure as “topped out”?

- How should potential for “back-sliding” performance on measures inform analyses and considerations?
- How should a lack of alternative measures in an important topic area inform analyses and considerations?

When a measure is determined to be “topped out”, what options should RAND consider for that measure in its analyses?

- Retired
- Display-Only
- Down-weighted