

RAND Medicare Advantage (MA) and Part D  
Contract Star Ratings Technical Expert Panel  
October 30th 2018 Meeting

**MEETING SUMMARY**

Cheryl L. Damberg and Susan M. Paddock



For more information on this publication, visit [www.rand.org/t/CF394](http://www.rand.org/t/CF394)

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2019 RAND Corporation

**RAND**® is a registered trademark.

#### Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit [www.rand.org/pubs/permissions](http://www.rand.org/pubs/permissions).

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

#### Support RAND

Make a tax-deductible charitable contribution at  
[www.rand.org/giving/contribute](http://www.rand.org/giving/contribute)

[www.rand.org](http://www.rand.org)



## RAND Technical Expert Panel Medicare Advantage (MA) and Part D Contract Star Ratings

October 30<sup>th</sup>, 2018

### MEETING ATTENDEES

#### Technical Expert Panel Members

NAME	POSITION
Arlene Ash, PhD	Professor and Division Chief for Biostatistics and Health Services Research in the Department of Quantitative Health Services, University of Massachusetts Medical School
Liza Assatourians, JD	Vice President of Federal Programs, America's Health Insurance Plans (AHIP)
Anne Burns, RPh	Vice President, Professional Affairs at the American Pharmacists Association (APhA)
Lindsey Copeland, JD	Policy Director, the Medicare Rights Center
Jennifer Eames Huff, MPH	Independent consultant; Senior Advisor, Pacific Business Group on Health
Eve Kerr, MD, MPH	Louis Newburgh Research Professor of Internal Medicine, University of Michigan Medical School, Director of the Ann Arbor VA Center for Clinical Management Research, and Director of the Michigan Program on Value Enhancement
Elisa Munthali, MPH	Senior Vice President of Quality Measurement, National Quality Forum
Amy Nguyen Howell, MD, MBA	Chief Medical Officer, America's Physician Groups (APG)
Deborah Paone, DrPH, MHSA	Performance Evaluation Lead for Quality Measurement, Social Determinants of Health, and Care Innovation, Special Needs Plans (SNP) Alliance
Ninez Ponce, MMP, PhD	Professor, University of California Los Angeles Fielding School of Public Health's Department of Health Policy and Management, Director of the Center for Global and Immigrant Health
Patrick Romano, MD, MPH	Professor of Medicine and Pediatrics, University of California Davis School of Medicine
Marissa Schlaifer, MS, RPh	Consultant, Pharmaceutical Care Management Association
Allyson Schwartz, MSS	President and CEO, the Better Medicare Alliance
Nadine Shehab, PharmD, MPH	Senior Scientist, Medication Safety Program in the Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention
Dolores Yanagihara, MPH	Vice President of Analytics & Performance Information, Integrated Healthcare Association (IHA)



**RAND Staff**

Cheryl Damberg, PhD (PI/Project Director)  
Susan Paddock, PhD (co-PI/Project Director)  
Marc Elliott, PhD  
Justin Timbie, PhD  
Rachel Reid, MD  
Andy Bogart, MS  
Jessica Phillips, MS



## MEETING SUMMARY

### Welcome and Introductions

- The RAND Project Director, Cheryl Damberg, began the meeting by welcoming attendees and taking attendance of Technical Expert Panel (TEP) members.
- Cheryl Damberg reviewed the agenda and meeting goals. During this second meeting of the TEP, RAND indicated it would provide updates on analyses conducted since the first TEP meeting and seek input on analyses RAND is conducting regarding: 1) the utility and feasibility of constructing and reporting Star Ratings at smaller geographic units, and 2) estimating contract performance within geographic areas.

### TEP Discussion

- The summary of the TEP's discussion is organized by the two main discussion topics for the October 30, 2018 meeting, which map to the two presentations within the PowerPoint slide deck that RAND used to guide the discussion:
  - The Utility and Feasibility of Constructing and Reporting Star Ratings at Smaller Geographic Units
  - Estimating Contract Performance within Geographic Regions

### Update on Threshold Analyses

- RAND co-Project Director, Susan Paddock, provided a quick update on progress in conducting analyses related to increasing the stability of measure thresholds used to assign stars. RAND noted that CMS is seeking comments on several options for creating greater stability in Star Rating thresholds, and the options for setting thresholds are contained in the proposed 2019 Part C & D rule, available in the Federal Register. The comment period ends December 31<sup>st</sup>, 2018.

### The Utility and Feasibility of Constructing and Reporting Star Ratings at Smaller Geographic Units

- Currently, CMS reports Star Ratings at the contract level. Because a single contract can cover multiple geographic areas and performance may vary across areas there is interest in considering reporting performance at smaller geographic units. During the first meeting of the TEP, the TEP commented that having performance information by geography would be beneficial to Medicare beneficiaries when making choices about their healthcare, but wanted to better understand the consequences of reporting at smaller geographic units.
- To explore the utility and feasibility of constructing and reporting Star Ratings at smaller geographic reporting units (GRUs), RAND conducted simulations using seven quality measures included in the Star Ratings, and applied three different definitions of GRUs: 1) census regions, 2) Medicare Advantage (MA) regions, and 3) states.
- RAND team member, Justin Timbie, presented preliminary findings from analyses to assess the feasibility of reporting at the GRU level. The analyses focused on understanding whether the measures would meet minimum sample sizes and reliability criteria when assessed at the level of the GRU. *A copy of this presentation can be found in the corresponding slide deck.* Questions posed to the TEP were as follows:



- ***Employer group health plans were included in the analysis. Should RAND reconsider this decision?***
  - Because a number of contracts have a small number of enrollees in many geographic areas, the TEP members expressed interest in RAND conducting a sensitivity analysis that would exclude employer health plans to determine if the results would change.
  - A TEP member expressed caution that excluding employer health plans might cause confusion or disruption in the market place.
  - If the findings do not support including employer health plans, CMS should consider developing different measurement approaches for different types of contracts, for example, different approaches for state-based MA plans versus employer health plans.
  - It was also noted that implementing smaller units of measurement would likely mean that a larger percentage of contracts will not meet sample size criteria for performance measures and thus be excluded from rating calculations.
- ***Should RAND limit measurement to beneficiaries in GRUs that meet the reliability/sample size criteria? In cases where GRUs do not meet the criteria, would reverting to contract-level reporting be appropriate?***
  - TEP members commented that it is reassuring that most beneficiaries were found to live in GRUs that met the measure minimum sample size criteria. One member commented that state-level reporting is important for dually-eligible beneficiaries because Medicaid is coordinated at the state level. This member expressed concern about reverting to contract-level reporting when small contracts, such as SNPs, do not meet sample size criteria– they are washed out by large numbers and have less meaningful information available.
- ***Given the large percentage of beneficiaries living in state-level GRUs with sufficient sample sizes and reliability, should RAND consider smaller geographic areas? Are there other methods to disaggregate contracts RAND should consider?***
  - One TEP member asked the group to take a step back and think about what CMS is trying to accomplish with potentially moving to finer units of reporting. This member noted it will be challenging to establish what a meaningful sub-region is. Are there other existing units that could be used? For example, could CMS consider using hospital referral region (HRR)? Additionally, CMS should calculate performance on measures at all available units of measurement when feasible. There will be places where measurement is not feasible and that is okay.
  - RAND responded that the purpose of the Star Ratings is to provide information to beneficiaries so they can make informed decisions when selecting contracts, and to determine Quality Bonus Payments for MA contracts. RAND commented that it will be difficult to produce scores for all entities at smaller geographic units of measurement.



- One member commented that a possible unintended consequence of geographic reporting is beneficiaries will not receive the same benefits/services as beneficiaries living in other areas because their region did not qualify for a Quality Bonus Payment. A contract should be responsible for all beneficiaries, living in all areas.
- ***Does it make sense to add more measures and/or simulate Star Ratings as the next step in better understanding the feasibility and implications of geographic reporting? What additional analyses should RAND consider?***
  - One suggestion was an analysis looking at what percent of contracts with a high percent of LIS/dual eligible beneficiaries are included. RAND agreed it would be helpful to look at the profile of contracts that did not meet inclusion criteria.
  - It was also suggested that RAND look at which GRUs lack concordance with contract-level Star Ratings.
  - TEP members agreed it would be helpful to simulate both measure results and bonus payment results.
  - There was interest expressed by the TEP to continue the geographic reporting discussion at future TEP meetings.

### **Estimating Contract Performance within Geographic Regions**

- RAND co-Project Director, Susan Paddock, presented on possible estimation approaches for geographic reporting that address some of the challenges discussed in the prior presentation, including small sample size per contract in a geographic region. A copy of this presentation can be found in the corresponding slide deck.
  - Three types of approaches were presented: 1) use data only from the GRU and time period of interest to estimate performance; 2) pool data for a GRU over time to obtain an estimate for the GRU; and 3) use shrinkage estimation, a statistical technique that borrows information across GRUs to increase stability of estimates, particularly for GRUs with estimates that are relatively high variance.
  - TEP members asked several questions about shrinkage estimation to fully understand that option.
  - One member asked a clarification question regarding approach 2a which borrows information across contracts - borrowing information from all contracts within the same geographic region means contracts would be influencing one another. The member expressed concern that under this approach, a small contract would not be able to compete in a region. RAND confirmed that approach 2a borrows information from all contracts within a region to increase stability of estimates for each contract within a region. A disadvantage of this approach is that small contracts are likely to have estimates pulled toward the average contract score.
  - There was a question about whether approach 2b uses a weighted average or an unweighted average across all contracts because with market heterogeneity, an unweighted average could be influenced by a few outliers.
  - One member expressed hesitation to produce scores when there is very little data available. We don't know much about how small versus large contracts perform because our definition of small is an artifact. Is there evidence available that looking at a large national contract's performance is valuable in determining what would happen within a subgroup of beneficiaries living within a geographic area?



- One member commented that regional effects may play a larger role than contract performance, and suggested examining regional effects by comparing contract performance within the same GRU.
- RAND agreed that a useful next step would be to examine features of contracts or regions such as contract size, sponsor types, etc.
- There was also a comment that some measures included in the Star Ratings are related to how the contract operates, such as call center and appeals measures, and should continue being reported at the contract level, while other measures are influenced by local influences such as providers and would be more amenable to a geographic reporting.
- ***Which contract characteristics are important to consider for inclusion in a shrinkage approach?***
  - It was noted that hospital characteristics and contract characteristics look different. Hospitals can't change some of their characteristics such as their location. Contracts could more easily redesign some of their characteristics to align with what would achieve the best shrinkage target. We do not want to include contract characteristics that can be easily manipulated.
  - One member of the TEP asked whether more clarity could be provided if quality scores were associated with providers in addition to geographic areas. One could envision contracts dropping low performing providers to achieve improved scores. Some contracts may already be doing this to improve their scores. Another member commented that contracts are required to meet provider requirements and they may not be able to drop a lower performing provider. RAND commented that RAND receives data at the person-level and contract level but not at the provider level, and is thus unable to do this analysis.
  - There was agreement it would be helpful to look at patient characteristics.
  - There was also a recommendation for RAND to look at nine-digit zip tract data to better understand neighborhood characteristics (e.g., poverty).
  - RAND confirmed it does have access to age, gender, race/ethnicity, disability status, and some census-block data.
  - Despite the specific suggestions about contract, neighborhood, and individual characteristics, TEP members agreed that they would like to develop a better understanding of the variation across contracts and GRUs before commenting on the best way to apply shrinkage estimation.
- One TEP member commented that contracts and providers will do all they can to adapt to a new measurement structure. As such, while there is interest in providing more detailed information to beneficiaries, there needs to be consideration of possible unintended consequences. The Star Ratings program has been very successful in driving quality improvement and measurement, and CMS should avoid veering away from that.
- RAND reiterated that the approaches shared during the meeting are preliminary and we are continuing to explore options.