RAND Technical Expert Panel Medicare Advantage (MA) and Part D Contract Star Ratings, November 19, 2020 Meeting

PRESENTATION

Cheryl L. Damberg, Anagha Tolpadi, and Melony Sorbero
Today’s meeting will focus on different approaches to creating health equity measures aimed at reducing disparities within the MA and PDP Star Ratings

- Approaches to increase awareness of disparities in performance
- Approaches to strengthen incentives to reduce/eliminate disparities
Approach 1

RAND has been exploring two approaches

1. Report measures of performance by beneficiary subgroups (i.e., stratified reporting)
2. Assess performance among those with social risk factors (SRF)

Increase awareness of disparities in performance

These are examples from a much broader array of approaches that could be used to increase awareness of disparities.
Stratified reporting of performance

• Stratified reporting refers to reporting performance on Star Rating measures by designated subgroups of plan members (e.g., duals vs. non-duals, disabled vs. non-disabled)

• An important first step to address disparities
  – Identifying subgroup performance differences is necessary for contracts to be able to take steps to reduce disparities
  – Sharing stratified performance information could help beneficiaries choose plans that best meet their specific needs (“for people like me”)

• Challenges—limited SRF data available and small sample sizes for some subgroups
CMS recently started requiring that MA contracts report stratified performance for 4 HEDIS measures

Measures reported to CMS
- Plan all-cause readmission
- Breast cancer screening
- Colorectal cancer screening
- Comprehensive diabetes care – eye exam

Contracts reported performance by 6 strata
- LIS/DE only
- Not LIS/DE and not disabled
- Disabled only
- Missing/Unknown
- LIS/DE and disabled
- Other

Note: CMS began stratified reporting with measurement year 2018
RAND explored the feasibility of stratified reporting

How many contracts would have sufficient data to allow reporting of stratified results to contracts and/or for public reporting?

- For 3 of the 4 measures*, we assessed separately for each of the 4 strata of interest
  - Number of contracts with at least 30 enrollees** in the denominator
  - Reliability of performance scores reported for each stratum (to assess the precision of the estimates for making comparisons)
  - Strata of interest were:
    1. LIS/DE only
    2. disabled only
    3. LIS/DE and disabled
    4. not LIS/DE and not disabled

*Analyses did not include plan all-cause readmission measure because specifications are being revised
**HEDIS specifications require 30 enrollees in the denominator for reporting of results
Most contracts had at least 30 enrollees in the denominator for each individual strata

*BREAST CANCER SCREENING*

Contracts with >= 30 in individual stratum denominator

- Contracts that received measure stars: 388
- LIS/DE only: 258
- Disabled only: 322
- Both LIS/DE and disabled: 283
- Not LIS/DE and not disabled: 334

* Among contracts receiving measure stars in 2020 Star Ratings
A large share of contracts do not have at least 30 enrollees in the denominator for LIS/DE only or both LIS/DE and disabled

**COLORECTAL CANCER SCREENING**

Contracts with >= 30 in individual stratum denominator

- Contracts that received measure stars: 402
- LIS/DE only: 118
- Disabled only: 319
- Both LIS/DE and disabled: 153
- Not LIS/DE and not disabled: 339

* Among contracts receiving measure stars in 2020 Star Ratings
Most contracts do not have at least 30 enrollees in the denominator for LIS/DE only, but do for other individual strata.

* Among contracts receiving measure stars in 2020 Star Ratings
The number of contracts that could report each of the four strata varies by measure*

* Among contracts receiving measure stars in 2020 Star Ratings
Alternative approach to meeting denominator criteria

• Combine beneficiaries across all 4 strata and then apply denominator criterion of 30 beneficiaries in contract

• Result:
  – Would increase the number of contracts for stratified reporting
  – Impact would vary by stratum and measure
    • Greatest impact for LIS/DE only and both LIS/DE and disabled strata for the colorectal cancer screening and diabetes care – eye exam measures
      – Number of contracts that could be reported more than doubles
    • Much smaller impact for disabled only and not LIS/DE and not disabled strata regardless of measure, and breast cancer screening regardless of stratum
      – Number of contracts that could be included increases by 11-25%
Reliability requirements ensure precise performance estimates

• High reliability enables differentiation between contracts’ performance
  – Performance estimates with high reliability reflect actual performance rather than measurement error

• Reliability is a ratio of “signal to noise”
  – A reliability of 0.7 can be interpreted as 70% of variation in performance is due to true variation and 30% is due to measurement error
  – Reliability increases with sample size

• CMS applies a reliability requirement in the current Star Ratings
  – For contracts with 500-1000 enrollees, a reliability of 0.70 is used for HEDIS clinical measures
Most contracts have high reliability for each stratum, but some contracts do not

Among contracts receiving measure stars in 2020 Star Ratings. Restricted to contracts meeting the denominator criteria of >=30.
Most contracts have high reliability for each stratum, but some contracts do not.

**COLORECTAL CANCER SCREENING**

* Among contracts receiving measure stars in 2020 Star Ratings. Restricted to contracts meeting the denominator criteria of $\geq 30$.  

- **LIS/DE only**: 96, 17, 5
- **Disabled only**: 206, 83, 30
- **Both LIS/DE and disabled**: 125, 22, 6
- **Not LIS/DE & not disabled**: 322, 8, 9

Legend:
- Reliability $\geq 0.7$
- Reliability 0.6 - <0.7
- Reliability < 0.6
Most contracts have high reliability for each stratum, but some contracts do not.

**DIABETES CARE - EYE EXAM**

* Among contracts receiving measure stars in 2020 Star Ratings. Restricted to contracts meeting the denominator criteria of >=30.
Discussion questions

STRATIFIED REPORTING

- Given the findings of our analyses, should CMS report contract-level stratified performance results (e.g., confidential to plans, display page, or on Medicare Plan Finder)?

- Should CMS pursue stratified reporting for additional HEDIS measures or non-HEDIS measures?

- Should CMS require a reliability of 0.7 or higher for stratified reporting, or are there scenarios where a lower reliability threshold is acceptable?

- Are there other analyses RAND should perform to inform decision making related to stratified reporting?
  - For example, analyses to inform whether and how to group SRF categories (see examples on next slide)?
#### Three possible SRF groupings

**Grouping 1:** Use 4 separate strata

<table>
<thead>
<tr>
<th>Both LIS/DE and disabled</th>
<th>Disabled-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIS/DE-only</td>
<td>Not LIS/DE &amp; not disabled</td>
</tr>
</tbody>
</table>

**Grouping 2:** Compare each SRF one at a time

<table>
<thead>
<tr>
<th>Both LIS/DE and disabled</th>
<th>Disabled-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIS/DE-only</td>
<td>Not LIS/DE &amp; not disabled</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Both LIS/DE and disabled</th>
<th>Disabled-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIS/DE-only</td>
<td>Not LIS/DE &amp; not disabled</td>
</tr>
</tbody>
</table>

**Grouping 3:** Compare LIS/DE or Disabled vs. neither SRF

<table>
<thead>
<tr>
<th>Both LIS/DE and disabled</th>
<th>Disabled-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIS/DE-only</td>
<td>Not LIS/DE &amp; not disabled</td>
</tr>
</tbody>
</table>
RAND has explored the construction of an SRF performance index (SRFPI)

- ASPE recommended the inclusion of health equity measures and/or domains in quality reporting programs and value-based purchasing (VBP)
- Goal is to incentivize contracts to perform well among socially at-risk beneficiaries
- The index summarizes performance among those with SRFs across multiple measures into a single score
- Uses only data for MA beneficiaries classified as LIS/DE or disabled
- SRF performance index could be used as basis for a reward added to the overall and summary Star Ratings or as a separate Star Ratings measure
Decision rules we applied in constructing the SRF performance index

• For a measure* to be included in SRF performance index:
  – At least 50% of contracts that received measure-level Star Rating met the SRF performance index sample size and reliability requirements for the measure

• For a contract to be eligible for SRF performance index:
  – At least 500 LIS/DE/disabled enrollees or at least 50% of enrollees are LIS/DE/disabled

• For a contract to receive SRF performance index:
  – Must have measure-level stars for at least 50% of measures in SRF performance index

*RAND considered the same measures that were adjusted for when constructing the 2021 CAI
Process for constructing SRF performance index

1. Measure data restricted to the LIS/DE/disabled beneficiaries
2. Measure denominator requirements applied and reliability calculated for each contract across LIS/DE/disabled
3. Measure thresholds and measure stars determined for eligible contracts
4. Created SRFPI for Overall, Part C and Part D resulting in an SRFPI score that ranges from 1 to 5
Simulated SRF performance index’s potential effect as a basis for reward

• Simulated two different performance targets to be eligible for reward
  – Contracts that received 4+ stars on index
  – Contracts that received 3.5+ stars on index

• Structure of reward:
  – Simulated a reward of 0.5 star
  – Contracts received reward if they met performance target on SRF performance index and had at least 5,000 LIS/DE/disabled enrollees or at least 20% LIS/DE/disabled enrollees
  – Reward was added to Overall and Summary Stars prior to final rounding
Most measures could be constructed for at least 50% of contracts.

- **MA-PD measures**: 3 measures met criteria, 16 did not meet criteria.
- **PDP measures**: 1 measure met criteria, 4 did not meet criteria.

* Purple and gray add up to the total # of measures considered for inclusion in SRFPI (i.e. measures included in 2021 CAI)
The SRF performance index could be calculated for most contracts

* Purple and gray add up to # non-Puerto Rico contracts receiving stars in 2020 Star Ratings
The average SRF performance index score is lower than the average unrounded Star Rating.

Note: Pre-application of SRF performance index reward.
Many contracts rewarded by the SRF performance index are high performing.
SRF performance index rewards are not concentrated in MA contracts with high percentage of LIS/DE or disabled beneficiaries.
No PDP contracts rewarded by the SRF performance index serve a high percentage of LIS/DE or disabled beneficiaries.
Next steps: Simulate including SRF performance index as Star measure

• Allows for a potential downside in addition to upside effect on Star Ratings

• Because performance on SRF performance index is lower than other measures in Star Ratings, including it would reduce most contracts’ Star Ratings

• Approaches we are considering to address this property include:
  – **Rescale the SRF performance index** by adding a constant to every contract’s index so the average star on the index is the same as the average star across other measures
  – Use the SRF performance index as an **improvement measure** (measure if performance among vulnerable populations is static or improves over time) with a hold harmless rule so that it does not negatively impact any high performing contract’s Star Ratings
Discussion questions
SOCIAL RISK FACTOR PERFORMANCE INDEX

Could inclusion of an SRF performance index or similar index/measure strengthen efforts to reduce disparities?

How might the SRF performance index be best used?

– To report information on the SRF performance index measure to contracts?
– To reward contracts (and what size reward is appropriate)?
– As measure in Star Ratings?

Are there specific analyses RAND should perform as it continues to advance the development of an SRF performance index?