Advising the Congress on Medicare issues

Factors affecting variation in Medicare Advantage plan star ratings

Carlos Zarabozo

September 10, 2015

MEDPAC

Presentation outline

- Review of Medicare Advantage star rating system and bonus provisions
- The issues prompting this analysis
- Commission findings and CMS findings
- Options for addressing the issue

The MA quality bonus program

- Plans rated using a 5-star rating system, with plans at 4 or more stars eligible for bonuses
  - Bonus is 5 percent add-on to benchmark (10 percent in some counties)
- Plan star level also determines rebate share when bid is below benchmark

Determining a plan's overall star rating

- 44 MA-PD quality measures include
  - Clinical process and outcomes
  - Patient experience
  - Contract administration
- Each measure receives a star rating
- Overall star rating is weighted average of 44 measures
  - Weighting system assigns most weight to improvement (weight of 5), outcomes (3), patient experience (1.5); least to process measures (1)

The issue

- Plans serving a primarily or exclusively low-income population attribute their poor performance in star ratings to the complex care needs and socioeconomic status of their enrollees
- An important issue because the MA program allows certain plans to exclusively serve Medicare/Medicaid dually eligible beneficiaries (D-SNPs)

Disability status as a factor

- Some plans exclusively serving Medicare-Medicaid dually eligible beneficiaries are able to achieve ratings of 4 stars or higher
- Why, if the population is so difficult to serve?
- Issue may be disability status of enrollees

Past Commission work: Disability status as a factor in plan-level results

- March 2015 report: Plans with higher shares of enrollees under the age of 65 (entitled to Medicare on the basis of disability) had lower overall star ratings
- Among D-SNPs, those enrolling only the aged had higher overall star ratings

First round of CMS findings: Low-income status as a factor

- CMS examined 19 of the 42 unique measures in the star rating system (excluding, for example, measures that are already case-mix-adjusted)
- For 6 of the 19 measures in MA, CMS found systematic differences between low-income enrollees and non-low-income enrollees that were statistically significant and of "practical significance"

CMS's original interim solution

- Discount (partially) the measures in which there are differences, thereby emphasizing measures for which there is a level playing field for cross-plan comparison
- CMS considered this an interim step, pending additional research
- Announced proposal in February 2015, but withdrawn after public comment—no change for 2016 stars

Most recent work

MedPAC

- Examined "person-level" results on a measure-by-measure basis for 36 measures
- Used quality measures together with demographic/risk data of enrollees (HMOs only)

CMS

- Continued analytic work looking at a variety of factors
- Added examination of disability status

CMS and MedPAC findings consistent

- Differences in several star measures based on low-income status (Medicare-Medicaid dual eligibility)
- Differences in several star measures based on disability status
  - Under age 65 (entitled to Medicare based on disability)
  - Age 65 but originally entitled based on disability
- Results not always worse for low-income or disabled enrollees

A question to answer

- Is poor performance due to the nature of the population being served, consistent with what plans have maintained, or
- Are the plans with high shares of certain populations lower quality plans?

Results within a large MA plan illustrate the population differences

Rates within a single plan: Percent of diabetics with poor control of HbA1c (blood sugar), 2012 [high rate of poor control = worse performance]

	Diabetics age >=65, original entitlement based on age	Age >=65, original entitlement under 65 via disability	Under age 65 (disabled)
Non-dual	7.5%	10.1%	16.9%
Dually eligible	10.8	13.7	19.2

Note: Dually eligible are eligible for Medicare and Medicaid. All differences statistically significant. Source: MedPAC analysis of HEDIS data for 2012 performance year and 2012 denominator file. DATA PRELIMINARY AND SUBJECT TO CHANGE

- **Rates among Medicare –Medicaid dually eligible higher (poorer performance) than among non-duals**
- **Bigger differences are differences between beneficiaries under age 65 and beneficiaries 65 or older**

The logic of CMS's original proposal

- The way MA is designed, certain plans have a high share of low-income individuals
- Plans are rated using a subset of available measures
- If the star subset included only those measures for which there is no evidence of potential bias against specific plans, there would not be an issue with the low-income population
- Reducing the star weights on measures showing potential bias partly gets at the issue

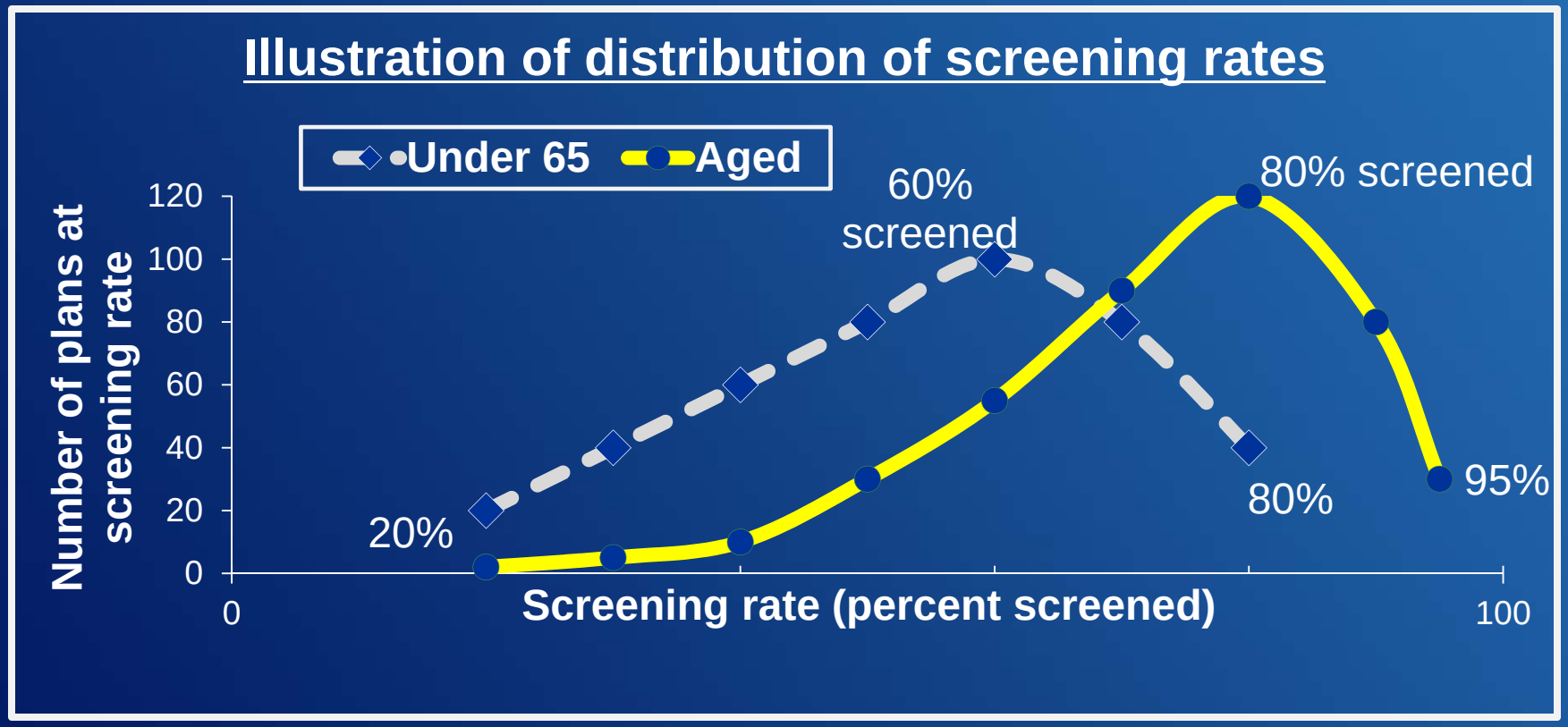
Possible approaches for a level playing field in determining bonus eligibility

- Peer grouping of plans based on composition of enrollment
- Star thresholds and performance levels determined by population groups

Alternative: Peer grouping of plans?

- Modeled after Commission approach to hospital readmissions penalties, where hospitals with high shares of low-income admissions have high readmission rates
  - Compare among peer groups for penalty purposes (e.g., deciles), based on hospital's share of low-income.
- Complex solution for MA across multiple measures, with population variation potentially different in each measure

An adjustment recognizing the differing distributions of rates by population groups



- In this illustration, across all plans the highest screening rate that any plan achieves for the aged population is 95 percent; for the under-65 (disabled), it is 80 percent.
- The most frequently achieved screening rate for the aged is 80 percent (120 plans); for the under-65 (disabled), 60 percent (100 plans).

MEDPAC

Establishing performance thresholds by population groups

Illustration of thresholds for the breast cancer screening measure with separate thresholds by population groups, aged and under 65

	Breast cancer screening rate cut-off points		
Performance threshold	**All beneficiaries combined**	**Under 65 (disabled)**	**Aged**
High (90th percentile)	>=83%	>=79%	>=85%
Medium high (75th)	78	74	80
Average (50th (median))	72	67	75
Below average (40th)	69	64	72
Low (30th)	66	62	70

Note: Percentile levels are not actual percentiles used for star cut points. Distribution of rates based on actual data for 2012 for HMOs with at least 411 beneficiaries in denominator for each category. DATA ARE PRELIMINARY AND SUBJECT TO CHANGE.

Different results with population grouping: The aged

Plan with only aged enrollees (age 65 or older), rate at 84 percent

	Breast cancer screening rate cut-off points		
Performance threshold	**All beneficiaries combined**	**Under 65 (disabled)**	**Aged**
	Rate of 84% = High		
High (90th percentile)	>=83%	>=79%	>=85%
			Rate of 84% = Medium high
Medium high (75th)	78	74	80

- **A plan that has only aged beneficiaries with a rate of 84 percent would be a high-performing plan when a combined threshold is used (current methodology) but would be medium-high if the two population groups had different thresholds.**

Different results with population grouping: The disabled

Plan with all enrollees under the age of 65 (disabled), rate of 76 percent

	Breast cancer screening rate cut-off points		
Performance threshold	All beneficiaries combined	Under 65 (disabled)	Aged
High (90th percentile)	>=83%	>=79%	>=85%
		Rate of 76%: Medium high	
Medium high (75th)	78	74	80
	Rate of 76%: Average		
Average (50th (median))	72	67	75

- **A plan with all beneficiaries under the age of 65, with a screening rate of 76 percent, would be average if there were not separate thresholds, but would be at the medium-high level with separate thresholds.**

MEDPAC

Result for plan with mixed population of aged and disabled

- Plan with mixed population of aged and disabled enrollees would have a weighted overall star rating
  - Use cut-points for the aged to determine the star rating for the aged enrollees in the plan (**Group A**).
  - Use cut points for disabled to determine the star rating for the disabled in the plan (**Group B**).
- Weighted average overall star rating is
  - Star rating for **Group A**, weighted by number of beneficiaries in **Group A**,

    PLUS
  - Star rating for **Group B**, weighted by number of beneficiaries in **Group B**.

Other methods

- Scaling results to proportionately raise or lower one set of results for comparability to another set of results
- Other possibilities

Magnitude of effect on overall star ratings

- In our initial simulation of effect on overall plan star ratings, only a few plans move from non-bonus status to bonus status
- May need to consider effect at overall star level in evaluating methods for addressing the issue

Commission discussion

- Questions on findings
- Discussion of options for addressing issue