

Kaiser Permanente Response to CMS Request for Information
**Data on Differences in Medicare Advantage (MA) and Part D Star Rating
Quality Measurements for Dual-Eligible versus Non-Dual-Eligible Enrollees**

Submitted by email to: PartCandDStarRatings@cms.hhs.gov

November 6, 2014

The Centers for Medicare & Medicaid Services (CMS) has requested research that addresses the question whether dual-eligible status (or “dual status”) causes lower performance on Star Ratings measure scores. Kaiser Permanente (KP) appreciates the opportunity to submit findings from an analysis of Star Ratings performance in our Southern California region comparing the performance of our dual eligible special needs plan (D-SNP) plan benefit package (PBP) on nine Star Ratings measures to that of non-SNP PBPs in the same region.

Given the length of the comment period, KP was unable to complete extensive analysis of all Star measures across all contracts. However, we are willing to continue to work with CMS to perform additional analyses, such as multivariate analysis of our D-SNP and non-D-SNP members’ measure scores against demographic, economic and health status characteristics to assist CMS in determining whether there is a causal relationship between dual status and Star measure scores.

Summary of Findings

The results of our analysis were mixed and do not conclusively demonstrate a difference in overall Star Ratings performance between duals and non-duals. Among our results, we found the following:

- For two HEDIS measures tested (Adult BMI Assessment and Glaucoma Screening), there were statistically significant differences in performance where the D-SNP performed better than the non-SNP population.
- For the three Part D medication adherence measures, the D-SNP performance was slightly worse than the non-SNP performance (but could not be tested at this time for statistical significance).
- For no measure we tested was the difference in the measure scores between D-SNP and non-SNP greater than approximately 6%.

Our full results are detailed in the tables and narrative further below, followed by a description of our data sources.

KP Study Design

CMS has requested “Analysis of the difference in measurement scores between dual and non-dual (or LIS and non-LIS) enrollees in the same contract and/or plan for all contracts under a parent organization for the Star Ratings measures.” CMS requests that the research findings demonstrate causality, not just an association.

In developing our analysis, we included the Star Ratings scores on select measures for 2012 and 2013 for the entire KP Southern California region, which represents the full Southern California

market within KP's California contract, H0524. We used our Southern California region D-SNP PBP as a proxy for "plans that enroll a disproportionate share of dual-eligible beneficiaries," since all members of D-SNPs are dually-eligible for Medicare and Medicaid. We note that KP's D-SNP members receive care from the same provider network as non-SNP members and, like all KP members, are free to choose their provider and switch at any time without compromising the goals of their SNP Model of Care. We segregated the Southern California D-SNP scores from the Southern California non-D-SNP scores and compared the results for 2012 and 2013.¹

In the analysis, we sought to address the following question: Does having large proportion of dually-eligible members cause a Medicare Advantage contract to have lower Medicare Star Ratings scores? This question can be deconstructed into the following questions:

1. In contracts with dual eligible members ("duals"), do duals have lower measure scores than non-duals?
2. Are the measure scores of non-duals in plans with dual eligible members lower than, equal to, or greater than measure scores in plans without dual eligible members?
3. Are the measure scores of non-duals in plans with a higher percentage of dual members lower than, equal to, or greater than the measure scores of non-duals in plans with a lower percentage of dual members?
4. Is the gap between dual and non-dual members in plans with a high percentage of dual members lower than, equal to or greater than the gap in plans with a low percentage of dual members?

To prove a causal relationship, and to determine the magnitude of the possible relationship, it is not enough to document that contracts with duals have lower Star measure scores than contracts without duals. Similarly, it is not enough to document that contracts with higher percentages of duals have lower measure scores than contracts with lower percentages of duals. It is possible that contracts with duals (or higher percentages of duals) have different overall levels of quality performance than contracts without duals (or with low percentages of duals).

For the two illustrative examples below, we assume that measure scores for duals are generally lower than measure scores for non-duals:

- If contracts with dual eligible members had higher levels of quality than contracts without dual eligible members, then the measure scores for their non-dual members would be higher on average than the measure scores of plans without duals. In this case, just comparing the average measure score of duals to non-duals across all plans would underestimate the lower scores resulting for the duals.
- In contrast, if contracts with duals had lower average quality, then just comparing the average measure score of duals to non-duals across all plans would overestimate the lower scores resulting for the duals.

KP is able to provide information to help address the first question listed above: Do our dual eligible members have lower scores than our non-dual members? We are providing a comparison

¹ Note that KP only has D-SNPs; we do not have C-SNPs or I-SNPs. Therefore, any reference in this document to KP's SNP refers to our D-SNP.

of the scores for nine measures broken out by D-SNP and non-D-SNP. If other plans provide these same break-outs then CMS may be better able to address all four questions underlying the causal relationship between duals status and lower Star Ratings. However, this analysis would not explain why duals' measure scores are different from non-duals' scores. To what extent might the differences be due to observable demographic, economic and health status characteristics? Knowing whether some of these characteristics were important predictors of the differences in dual vs. non-dual measure scores could provide important information to guide targeting efforts to improve the quality performance for dual eligible members.

We note that this analysis has the following limitation: the scores of the non-SNP population are influenced by the inclusion of members who are likely dually-eligible (i.e., they meet the low-income requirements to qualify for Medi-Cal) but have not enrolled in the D-SNP product.

Measure Selection

Measures were selected based on several criteria. We included only measures that have been included in the Medicare Star Ratings for a minimum of two years. We included measures whose outcomes were derived from large enough sample sizes to extract from the data the non-SNP performance for comparison to corresponding SNP performance and Combined (i.e., SNP + non-SNP) performance. Finally, we wanted to produce data for a satisfactory number of measures that would provide an adequate cross-section of data in a region with a large population. We were able to assemble an appropriate array of measures in KP's Southern California region, which is the largest market by population within its California Medicare contract (H0524). Five HEDIS and four Part D Pharmacy measures were selected as satisfying these criteria.

Analysis/Results

We initially sought to verify whether SNP member scores are lower than the corresponding non-SNP scores on the same measures. To that end, we have compared the two rates by computing the difference between the SNP and non-SNP rates in the following table.²

² In the tables, "SCA" refers to KP's Southern California region.

Δ SNP vs non-SNP Part C & D Measure Scores

		SCA	
		2012	2013
HEDIS	Measure		
	Cardiovascular Care - Cholesterol Screening	(0.23%)	(1.41%)*
	Adult BMI Assessment	0.23%*	0.35%*
	Osteoporosis Management In Women Who Had A Fracture	3.97%	0.37%
	Rheumatoid Arthritis Management	0.21%	(0.75%)
	Glaucoma Screening in Older Adults	2.23%*	3.52%*
Pharmacy	Diabetes Treatment	0.23%	0.17%
	Medication Adherence for Diabetes Medications	(1.73%)	(2.13%)
	Medication Adherence for Hypertension (RAS)	(3.35%)	(3.84%)
	Medication Adherence for Cholesterol (Statins)	(5.98%)	(6.19%)

Notes:

(1) Positive (+) Δ indicates SNP > non-SNP measure score for same measure and year.

(2) Negative (-) Δ indicates SNP < non-SNP measure score for same measure and year.

*Score differences significant at p<0.05 level. Due to data limitations, significance testing was not done for Pharmacy measures.

Differences between SNP and non-SNP measure scores in the table above are mixed and inconclusive as to whether SNP members score lower than their non-SNP counterparts. Across the two years reported, nine differences were positive. SNP performance on six of these measures was within 0.17% to 0.37% of the non-SNP performance, suggesting that, in these instances, there are not meaningful differences between scores for SNP members and non-SNP members. Additionally, two out of nine negative variances also measured less than 1%, which could also suggest that SNP member performance matches that of non-SNP members. Only for three measures – the Part D medication adherence measures – was there consistent evidence of lower SNP performance.

We tested the statistical significance of the score differences for the HEDIS measures; relevant data to test the statistical significance for the pharmacy measures were not available at the time of this writing. The SNP and non-SNP scores for Osteoporosis Management and Rheumatoid Arthritis Management are not statistically significantly different. The score differences for Glaucoma Screening in Older Adults are statistically significant, indicating that performance for SNP member is somewhat higher than for non-SNP members. SNP members also have higher scores than non-SNP members on Adult BMI Assessment, but the difference is very small and not clinically meaningful. The only other statistically significant difference is for Cardiovascular Care – Cholesterol Screening in 2013, in which SNP members scored slightly lower than non-SNP members.

Understanding whether the presence of SNP members results in lower overall scores does not lie solely in validating that SNP scores are simply lower than non-SNP scores. The proportion of SNP versus non-SNP data in the overall results is also important.

SNP % of Total Beneficiaries in Measure Sample

		Measure	SCA	
			2012	2013
HEDIS		Cardiovascular Care - Cholesterol Screening	6.10%	6.05%
		Adult BMI Assessment	6.12%	6.03%
		Osteoporosis Management In Women Who Had A Fracture	6.92%	6.68%
		Rheumatoid Arthritis Management	8.68%	8.88%
		Glaucoma Screening in Older Adults	5.15%	5.21%
Pharmacy		Diabetes Treatment	11.58%	11.74%
		Medication Adherence for Diabetes Medications	10.80%	10.81%
		Medication Adherence for Hypertension (RAS)	8.54%	8.57%
		Medication Adherence for Cholesterol (Statins)	8.07%	8.06%

For Part C HEDIS measures, we measured the percentage of the combined measure denominator composed of samples from the SNP population. Across these five measures in both years, SNP members comprised approximately 5-9% of the beneficiaries for the measure sample. For the Part D Pharmacy measures, SNP beneficiaries comprised roughly 8-12% of all beneficiaries contributing to the overall measure sample. As both tables indicate, the significance of SNP member data in measure samples is relatively small compared to the non-SNP data.

Lastly, we measured the impact of SNP scores on the overall measure scores by computing the difference between overall and non-SNP measure scores:

Δ Combined vs non-SNP Part C & D Measure Scores

		Measure	SCA	
			2012	2013
HEDIS		Cardiovascular Care - Cholesterol Screening	(0.01%)	(0.09%)
		Adult BMI Assessment	0.01%	0.02%
		Osteoporosis Management In Women Who Had A Fracture	0.27%	0.02%
		Rheumatoid Arthritis Management	0.02%	(0.07%)
		Glaucoma Screening in Older Adults	0.12%	0.18%
Pharmacy		Diabetes Treatment	0.03%	0.02%
		Medication Adherence for Diabetes Medications	(0.21%)	(0.26%)
		Medication Adherence for Hypertension (RAS)	(0.31%)	(0.36%)
		Medication Adherence for Cholesterol (Statins)	(0.52%)	(0.54%)

Notes:

- (1) Positive (+) Δ indicates Combined > non-SNP measure score for same measure and year.
- (2) Negative (-) Δ indicates Combined < non-SNP measure score for same measure and year.

In the table above, a negative value indicates that the overall score was *less* than the non-SNP score, which further means that the SNP rate contributed to the decrease in the non-SNP rate. It follows that negative values in the prior ‘SNP vs non-SNP’ table are associated with negative values in the above table. However, as we saw earlier, variances between SNP and non-SNP scores were fairly narrow in most cases. The smaller proportion of SNP data in the overall measure sample mitigates the effect of the SNP rate even further in the overall Medicare rate. Given these two factors, the impact of SNP scores on the overall measure rates for Part C HEDIS and Part D Pharmacy measures remains relatively small in KP’s samples.

Data Sources & Data Integrity

HEDIS measure data were sourced from submissions to NCQA, as well as from data submitted by KP’s Clinical Analysis department for this study. Annual SNP-only rates were generated for two measures – Glaucoma Screening and Osteoporosis Management – using an administrative collection methodology. The remaining three HEDIS measures’ SNP-only rates were derived solely for this study from administrative level data KP’s Clinical Analysis department provided.

Part D Pharmacy measure scores were supplied by KP’s pharmacy analytical group. Rates reflect submitted prescription drug events (PDEs) and, therefore, may not reflect actual adjudicated rates reported by Accumen.

KP sought to incorporate data from an additional seven HEDIS measures, but variations in the measure collection methodology precluded their inclusion at this time. For example, specifications for Controlling High Blood Pressure strictly apply a “hybrid” collection methodology that designates use of a random sample for chart review. The hybrid collection methodology utilizes a sample of the respective data sets to measure SNP and overall scores independent of one another. Since the samples for these two populations may be of approximately equal size under hybrid collection, in the absence of administrative data, the proportion of the SNP population in the overall data cannot be determined from these samples. As a result we cannot segregate out the non-SNP scores.

Similarly, six other measures KP sought to include allow either administrative or hybrid data collection methods to be applied. In the case of five Comprehensive Diabetes Care measures (Cholesterol Screening, Eye Exam, Kidney Disease Monitoring, Blood Sugar Controlled, Cholesterol Controlled), KP’s Clinical Analysis department provided administrative collection data for SNP rates but hybrid collection data for overall scores for calendar year 2012. Since administrative data collection produced a SNP data set that was larger than the hybrid collection methodology used for overall scores, we could not segment out the non-SNP population from the overall measure data to compute a comparative rate. The last of the six measures, Colorectal Cancer Screening, used administrative data collection for both SNP and overall scores for 2012 rates and hybrid data collection for 2013. Therefore, KP was unable to produce results for these measures across two consecutive years and was compelled to exclude these measures.