

Examining the Potential Effects of Socioeconomic Factors on Star Ratings*



Elizabeth Goldstein
CMS, Center for Medicare

Cheryl Damberg, PhD
RAND

September 8, 2015

**The research presented is sponsored by CMS under contract HHSM-500-2013-00283G and performed by the RAND Corporation. The RAND Team included the work of Melony Sorbero, Ann Haas, Cheryl Damberg, Marc Elliott, and Susan Paddock.*

Issue to be Examined

Multiple Medicare Advantage (MA) organizations and Prescription Drug Plan (PDP) sponsors believe that enrollment of a high percentage of dual eligible (DE) enrollees and/or enrollees who receive a low income subsidy (LIS) disadvantages their plan's ability to achieve high MA or Part D Star Ratings.

Goal of Research

Provide the scientific evidence as to whether MA or Part D sponsors that enroll a disproportionate number of vulnerable beneficiaries are systematically disadvantaged by the Star Ratings.

Parallel analyses are being conducted to determine if modifications are needed for the payment risk adjustment models.

Star Ratings Policy

Star Ratings

- MA and Part D sponsors have responded to the Star Ratings program because it employs a solid, reliable methodology.
- The Star Ratings Program has led to quality improvements.
- CMS continuously reviews and makes improvements to the methodology to continue to incentivize plans and provide information that is a true reflection of the performance and experience of the enrollees.

Policy Goals

CMS is firmly committed to building the foundation for a solution that appropriately addresses the issue at hand and aligns with our policy goals:

- Recognizing the challenges of serving vulnerable populations and providing incentives for a continued focus for improving health care for these important groups;
- Proposing adjustments that reflect the actual magnitude of the differences observed in the data;
- Providing valid quality ratings to facilitate consumer choice;
- Providing incentives for MA and Part D quality improvement; and
- Recognizing the need for options that are both transparent and feasible for the plans and CMS to implement.

Background

Previous Findings

- Analyses by the Medicare Payment Advisory Commission* (MedPAC) suggest that Disability could be driving the differences observed in performance in the Star Ratings rather than low income status.
- The distribution of Overall Star Ratings by percentage of LIS/DE shows that some contracts with a higher percentage of LIS/DE enrollees can achieve high Star Ratings.
- CMS research included both examination of an LIS/DE effect and/or Disability effect.
 - These slides focus primarily on MA.

* Medicare Payment Advisory Commission, *Report to the Congress: Medicare Payment Policy* (Washington: MedPAC, March, 2015), is publically available at http://www.medpac.gov/documents/reports/mar2015_entirereport_revised.pdf?sfvrsn=0

Previous Findings (continued)

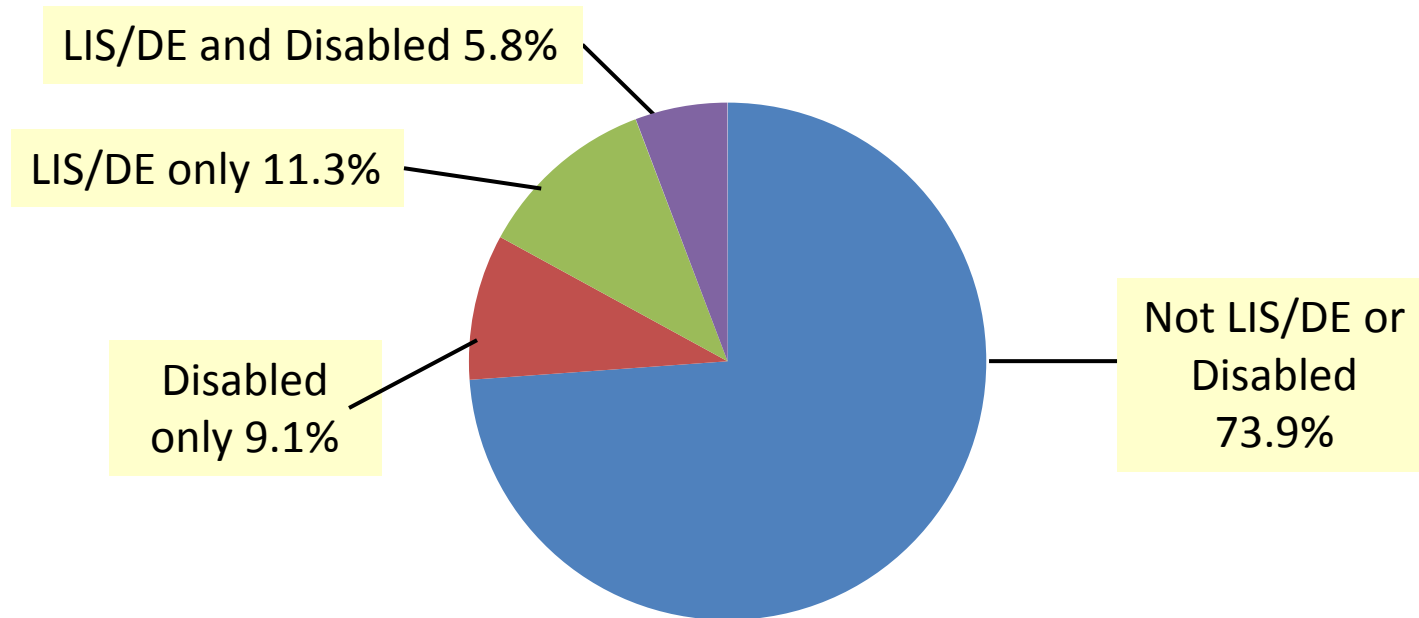
- The National Quality Forum (NQF) recommends assessing performance measures to determine if they should be adjusted for socio-demographic factors, including socio-economic status (SES), particularly when used for accountability and payment applications.*
 - Medicare Consumer Assessment of Healthcare Providers and Systems (CAHPS) and Health Outcomes Survey (HOS) are already adjusted for SES characteristics of patients, including education, LIS, DE, and health status.

* *National Quality Forum. Risk adjustment for socioeconomic status or other sociodemographic factors. Washington, DC: National Quality Forum; 2014*

Percentage of MA Beneficiaries by Status

More than one-quarter of MA Beneficiaries are LIS/DE, Disabled*, or both LIS/DE and Disabled.

Percent of all MA Beneficiaries**



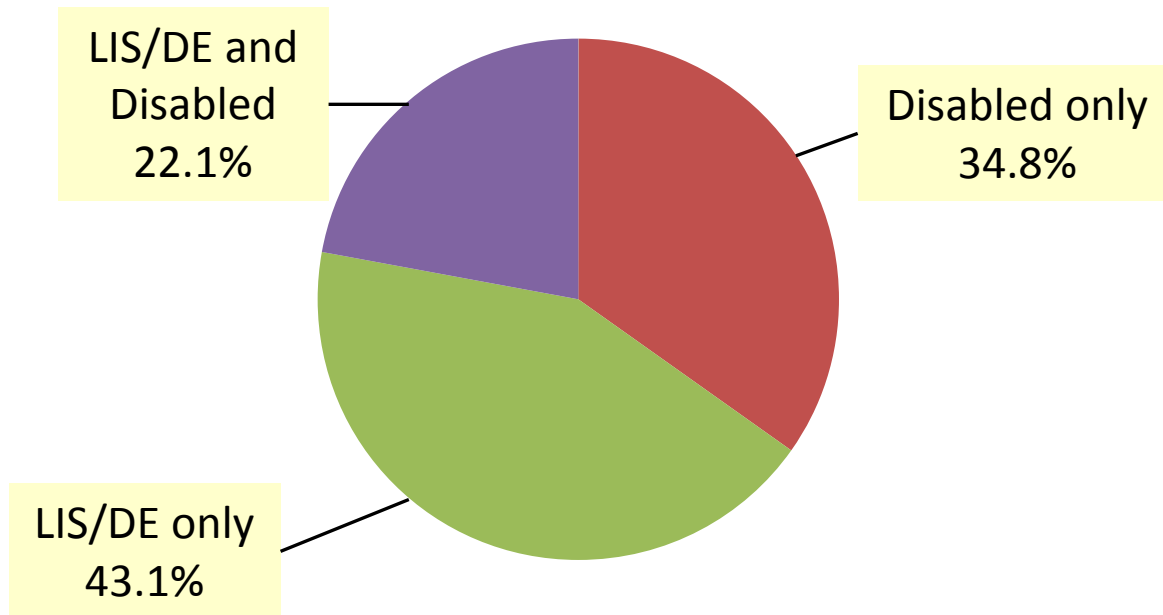
* Disability is based on the reason for entitlement as of December, 2012.

** Data source: Denominator file, 2012

Percentage of MA Beneficiaries by Status

There is overlap in the LIS/DE and Disability* status groups:

Percent of all LIS/DE or Disabled MA Beneficiaries**

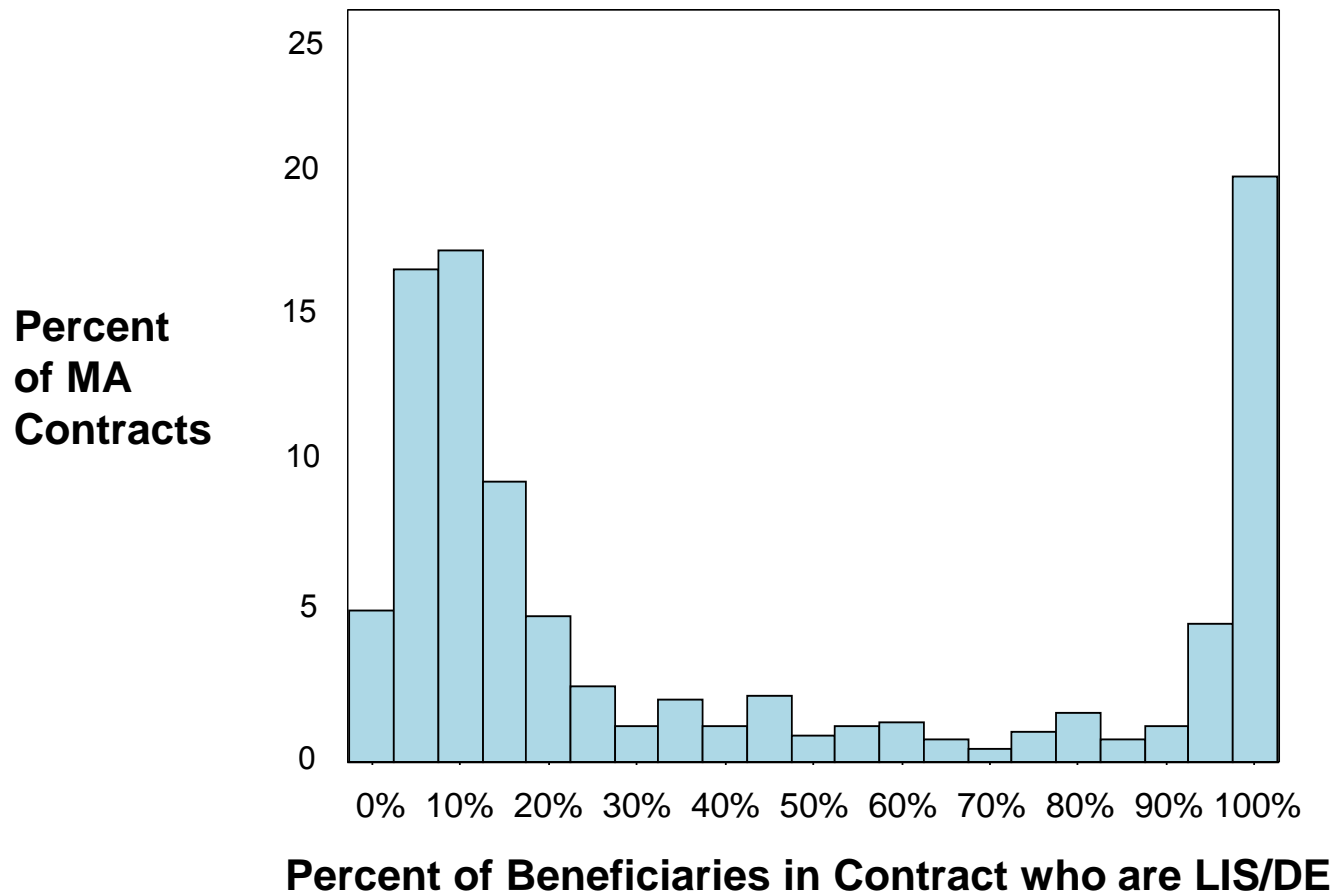


* Disability is based on the reason for entitlement as of December, 2012.

** Data source: Denominator file, 2012

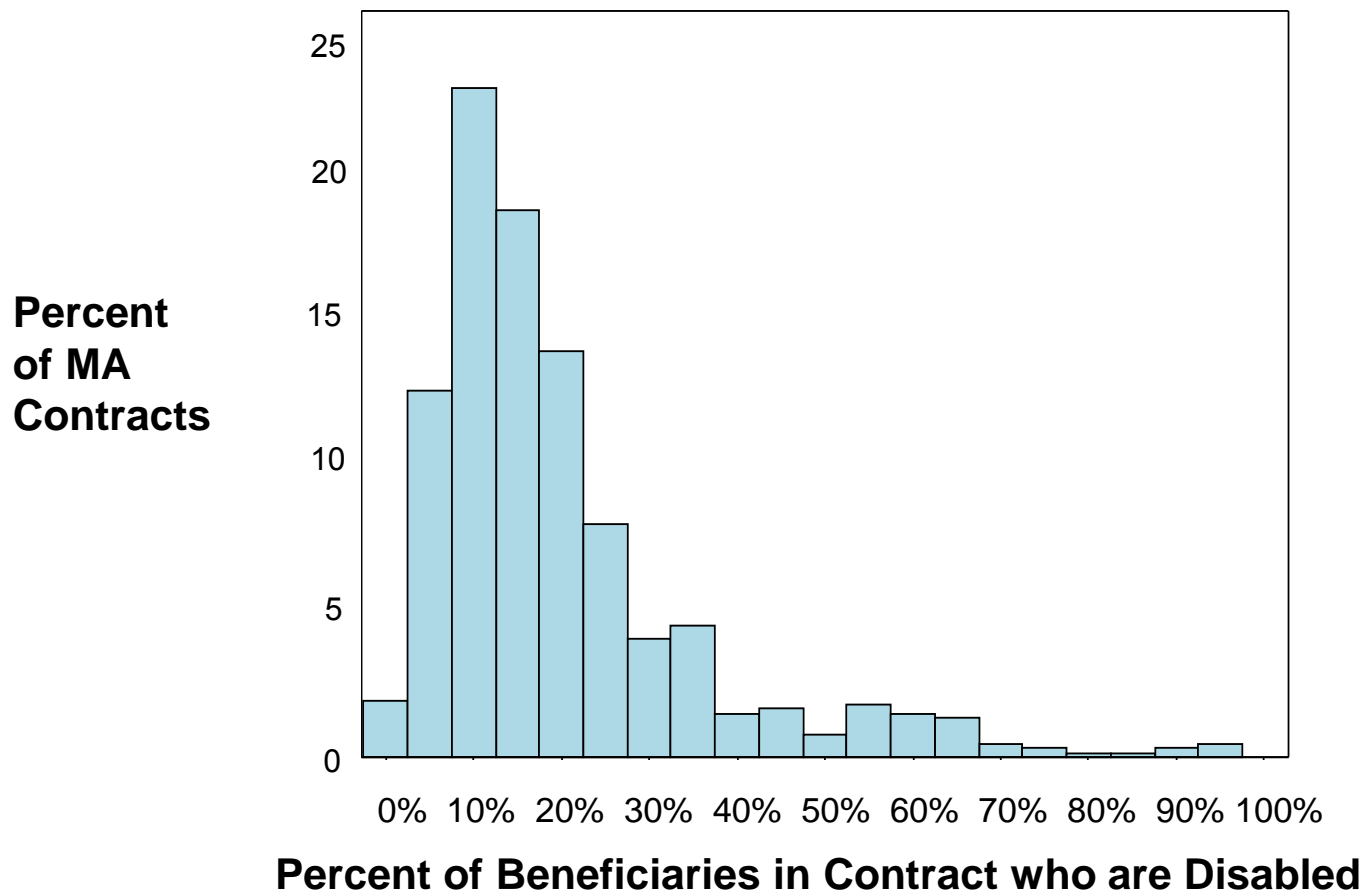
Distribution of MA Contracts based on Percentage of LIS/DE Beneficiaries

Contracts tend to either have a very high percentage (greater than 95%) of LIS/DE beneficiaries or a very low percentage (less than 20%) of LIS/DE beneficiaries.

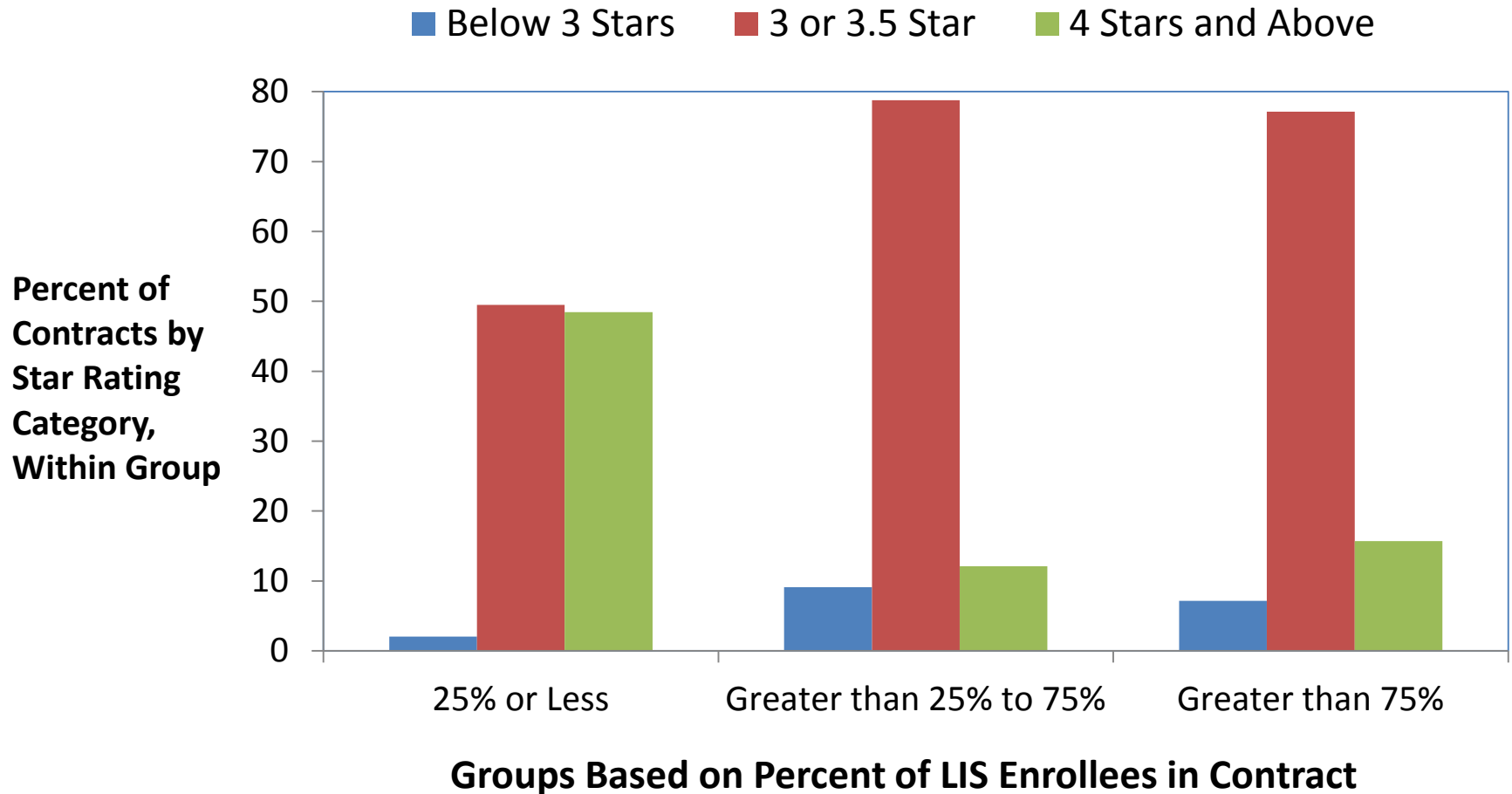


Distribution of MA Contracts based on Percentage of Disabled Beneficiaries

The distribution of MA contracts by percentage of Disabled beneficiaries is skewed. The majority of contracts have less than 25% of their enrollees classified as Disabled.



Distribution of Overall Star Ratings Grouped by Percentage of LIS Beneficiaries



Types of Disparity

When examining differences in the outcomes between the LIS/DE and/or Disabled groups and the non-LIS/DE and/or non-Disabled groups, the overall or total differences can be decomposed into between and within-contract differences (disparities). (*Refer to Appendix A for further details.*)

Between-Contract Disparities:

These differences represent true differences in quality between plans and are not appropriate for adjustment.

Within-Contract Disparities:

These are differences between subgroups for a particular contract.
These differences may be appropriate for adjustment.

Disparities in Care

- Many prior studies focused on the total difference (both between and within-contract differences).
- Our research focused on within-contract disparities, while controlling for between-contract disparities.
- Between-contract disparities: Lower SES beneficiaries are often enrolled in lower-performing contracts.
- Within-contract disparities: Lower SES beneficiaries may receive worse care within the same contract.

If there are no within-contract disparities, it may mean the contract successfully closed the quality gap or performance lag for all enrollees. It is also possible that contracts with no gap have below average quality performance for all enrollees.

Case Mix Adjustment (CMA)

One way to address within-contract disparities is through CMA.

The goal of CMA is to make contract comparisons more equitable. In other words, making adjustments as if each contract is serving the same mix of beneficiaries.

CMA encourages plans to serve vulnerable beneficiaries because it levels the playing field, so that all contracts have equal ability to achieve a high Star Rating, regardless of their percentage of vulnerable beneficiaries.

CMA (continued)

- A CMS priority is to use consensus-based measures.
 - The developers for Star Ratings measures include:
 - Healthcare Effectiveness Data and Information Set (HEDIS) developed by the National Committee for Quality Assurance (NCQA)
 - Pharmacy Quality Alliance (PQA)
 - Measures endorsed by the National Quality Forum (NQF).
- Measure specifications can only be changed by the measure steward (the owner and developer of the measure).
- Measure development and endorsement is a multi-year process.
- While measures are being reviewed by the measure developers, other interim calculation options could be used to adjust Star Ratings.

Measures Examined

Selecting Measures for Analysis

- A measure was *excluded* from analysis if:
 - The measure was already case-mix adjusted for socio-economic status (i.e., CAHPS and HOS measures).
 - The focus of the measurement was not a beneficiary-level issue, but rather a plan-level issue (i.e., appeals, call center, Part D price accuracy).
 - The measure was scheduled to be retired or revised.
 - The measure was applicable to only Special Needs Plans (SNPs) (i.e., SNP Care Management, Care for Older Adults measures).
- A total of 16 clinical measures were included in this analysis.

Clinical Measures Examined

Used 2014 Star Ratings data (2012 measurement year patient-level data)

HEDIS Measures

Adult BMI Assessment

Rheumatoid Arthritis Management

Breast Cancer Screening

Controlling High Blood Pressure

Diabetes Care – Blood Sugar Controlled

Diabetes Care – Eye Exam

Diabetes Care – Kidney Disease Monitoring

Colorectal Cancer Screening

Osteoporosis Management in Women who had a Fracture

Plan All-Cause Readmissions

Annual Flu Vaccine (Assessed through CAHPS, patient self-reported)

HOS/HEDIS Measures

Monitoring Physical Activity

Reducing the Risk of Falling

Pharmacy Quality Alliance Measures

Medication Adherence for Diabetes Medications

Medication Adherence for Hypertension

Medication Adherence for Cholesterol

Associations Between Performance and Vulnerable Beneficiaries (i.e., LIS/DE and/or Disability)

Associations Between Performance and LIS/DE and/or Disability*

- To determine whether the selected Star Ratings measures are sensitive to the composition of the enrollees in a plan, we examined the relationship between the performance and LIS/DE and/or Disability status.
- Odds Ratios are used to report the association between performance and LIS/DE and/or Disability status.

Note: Caution must be used when interpreting Odds Ratios. Large sample sizes can lead to statistically significant findings that are not necessarily meaningful. For example, an Odds Ratio of 0.99 may be statistically significant, but in a practical sense, does not lead to the interpretation of a difference in outcomes between the two groups.

* Refer to the Appendix A for further details about interpreting Odds Ratios and Appendix B for additional analyses.

Likelihood of Receiving Recommended Care or Outcomes

HEDIS Measure (MA Contracts)	LIS/DE Adjustment Odds Ratio	Disability Adjustment Odds Ratio
Adult BMI Assessment	1.11***	0.93***
Rheumatoid Arthritis Management	0.85***	1.17***
Breast Cancer Screening	0.69***	0.72***
Controlling High Blood Pressure	0.99	1.02
Diabetes Care – Blood Sugar Controlled	0.68***	0.63***
Diabetes Care – Eye Exam	0.93***	0.68***
Diabetes Care – Kidney Disease Monitoring	0.93***	0.69***
Colorectal Cancer Screening	0.87***	0.47***
Osteoporosis Management in Women who had a Fracture	0.71***	0.56***
Plan All-Cause Readmissions [#]	0.87***	N/A ^{&}
Annual Flu Vaccine	0.85***	0.72***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

[#] Measure is reverse-coded to make interpretation of Odds Ratio the same as other measures.

* Significant at p<0.05 ** Significant at p<0.01 *** Significant at p<0.001

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being LIS/DE or Disabled.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being LIS/DE or Disabled.

Black Odds Ratio indicates no significant effect. (No * indicator.)

[&] Not further adjusted for Disability. Readmissions is adjusted for factors that might be part of a person's reason for Disability.

Likelihood of Receiving Recommended Care or Outcomes

HOS and PDE Measure (MA Contracts)	LIS/DE Adjustment Odds Ratio	Disability Adjustment Odds Ratio
Monitoring Physical Activity	0.98	1.34***
Reducing the Risk of Falling	1.67***	1.32***
Medication Adherence for Diabetes Medications [^]	0.94***	0.75***
Medication Adherence for Hypertension [^]	0.86***	0.72***
Medication Adherence for Cholesterol [^]	0.94***	0.79***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

* Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being LIS/DE or Disabled.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being LIS/DE or Disabled.

Black Odds Ratio indicates no significant effect. (No * indicator.)

[^]The sample sizes for PDE were very large, so very small differences become statistically significant

Summary of Associations Between Performance and LIS/DE or Disability

Type of Association	LIS/DE	Disability
Positive	2	3
No effect (not significant)	2	1
Negative	12	11
Total Number of Measures	16	15*

** Readmissions is excluded as it is already adjusted for several factors that could determine Disability status.*

The findings were somewhat mixed:

- Most associations were negative. This means that vulnerable beneficiaries were less likely than non-vulnerable beneficiaries to receive the recommended care or outcome.
- For a small subset of measures, the association was positive. In other words, vulnerable beneficiaries were more likely to receive the recommended care or outcome. There were more measures that demonstrated a positive association with Disability as compared to LIS/DE.
- There were some measures for which there was no significant association between LIS/DE or Disability and receiving the recommended care or outcome.

Summary of Additional Associations Examined*

- Other Patient Characteristics
 - LIS/DE effect not sensitive to inclusion of other patient characteristics (age, gender, HCC, end-stage renal disease).
 - LIS/DE associations were not as strong when controlling for other patient characteristics.
- Census-Block Variables
 - Little improvement in explanatory power of model when census block SES factors are included.
 - Not sensitive to inclusion of census block group-level SES measures (education, income/poverty).

** Further details of the additional associations study are not provided because the findings support the focus to be on potential adjustment for LIS/DE and/or Disability.*

These analyses included controls for between-contract differences, LIS/DE, Disability status, and either other patient characteristics or census block characteristics.

Examining Variation of Within-Contract Disparities

Examining Variation in Within-Contract Disparities

Question to be addressed:

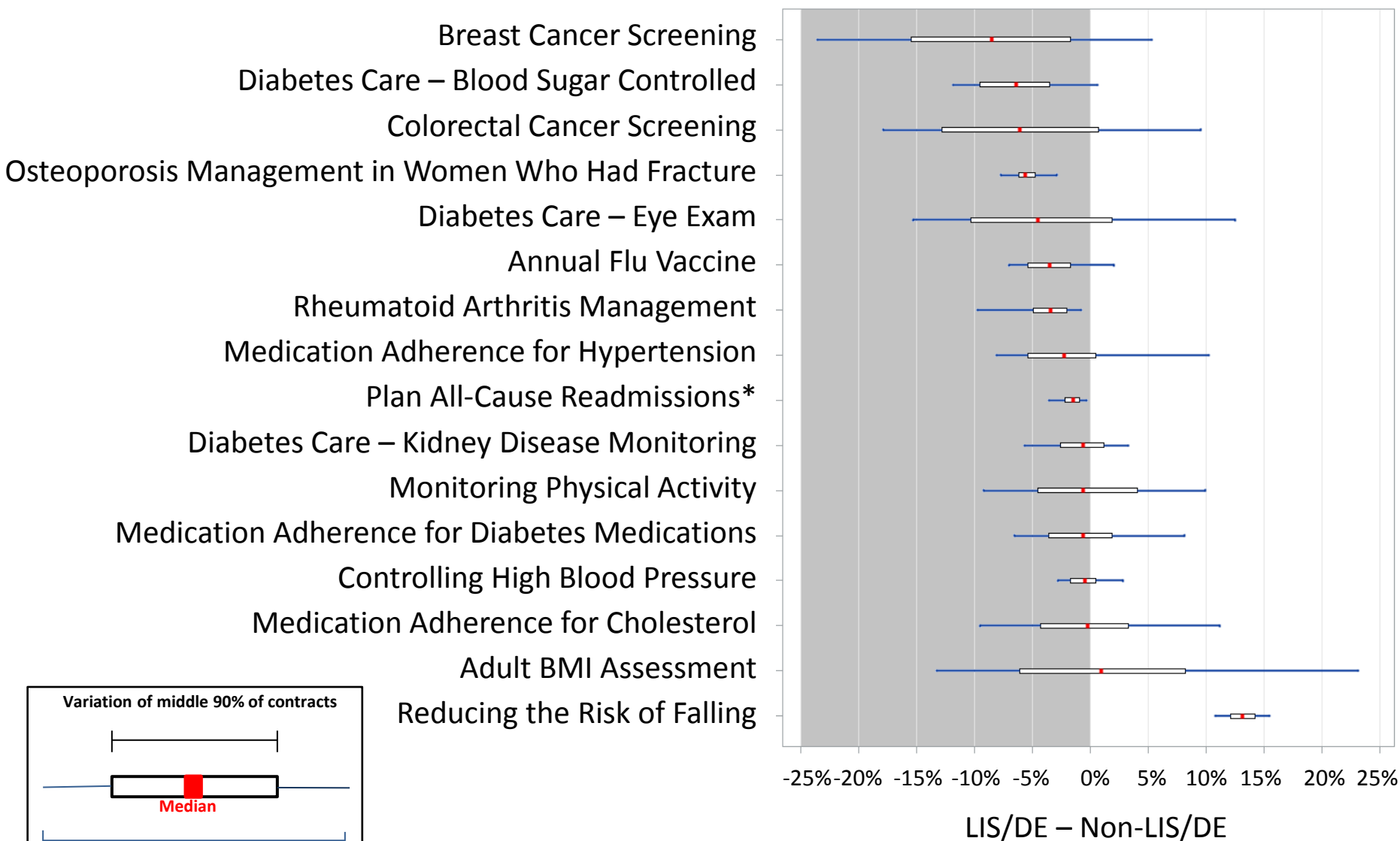
- Is there a difference (“disparity”) in performance for enrollees who are LIS/DE versus non-LIS/DE within contracts?
 - *Does this disparity vary by contract?*

Method for Examination

- For each contract, the proportion of beneficiaries receiving the measured clinical process or outcome for LIS/DE and non-LIS/DE beneficiaries was estimated separately, and the difference between the LIS/DE and non-LIS/DE performance rates per contract was calculated.
 - *A parallel analysis was performed for Disability.*

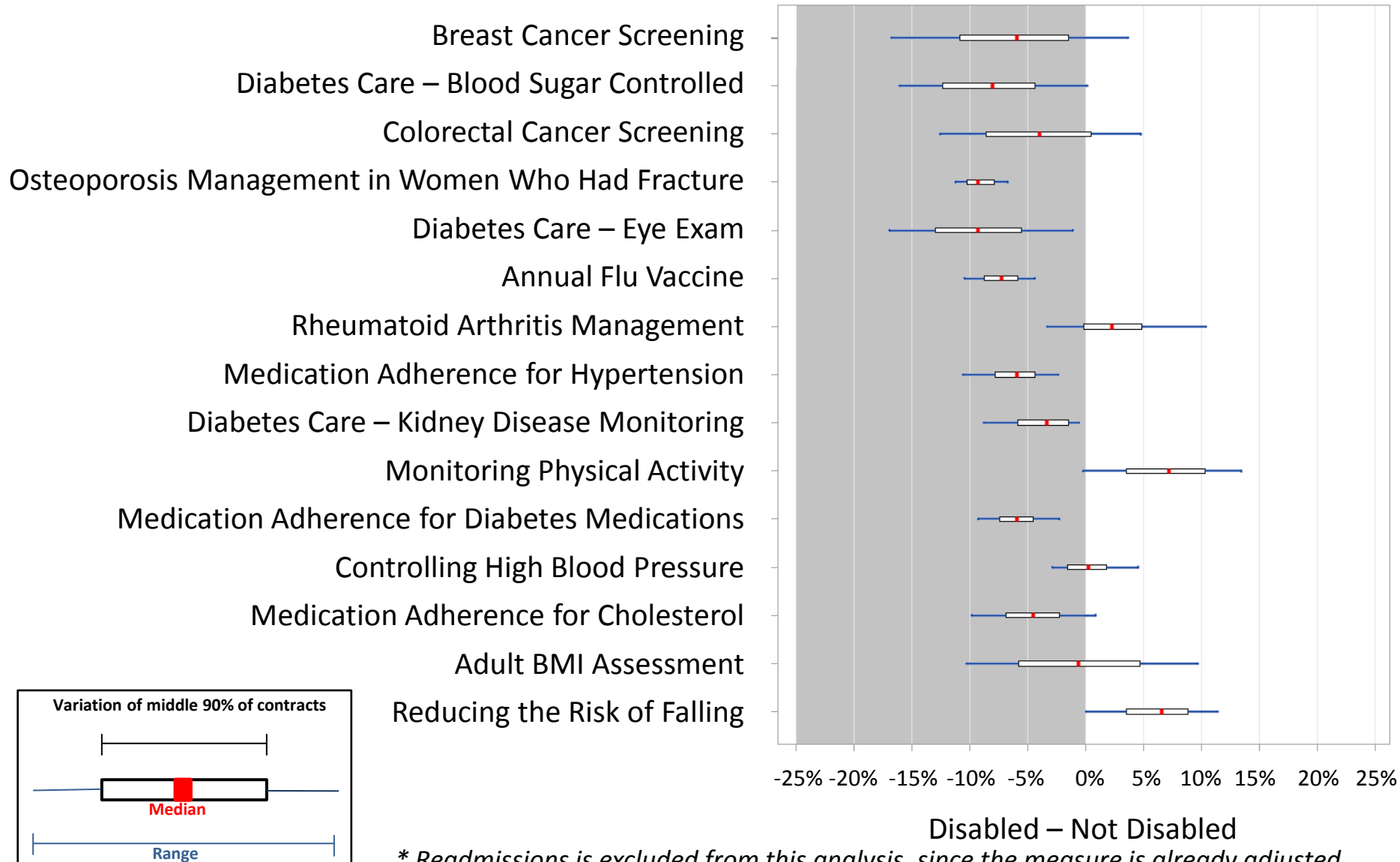
**Model accounts for differences in contract quality for LIS/DE (or Disabled) and non-LIS/DE (or non-Disabled) beneficiaries.*

Variation Across MA Contracts in Within-Contract LIS/DE Disparity



*Readmissions was reverse coded to be in the same direction as the other measures in the graphic.

Variation Across MA Contracts in Within-Contract Disability Disparity*



* Readmissions is excluded from this analysis, since the measure is already adjusted for factors associated with Disability status.

Summary of Examination of the Variation of Within-Contract Differences

- There is evidence of differences in performance between LIS/DE and non-LIS/DE, and Disabled and non-Disabled within a contract.
 - *The overall variability of the within-contract disparities differed across measures.*
 - *The median within-contract disparity indicated a small LIS/DE or Disability effect.*
 - *For some measures, the LIS/Dual or Disabled group outperformed their counterparts.*
- If focusing strictly on an LIS/DE effect, the results are mixed. The majority of the measures examined had a small, negative within-contract difference or no difference. This means for some measures (negative within-contract difference) the LIS/DE beneficiaries have worse outcomes as compared to the non-LIS, but for other measures (no difference) the LIS/DE have similar outcomes than the non-LIS/DE. For one measure, the LIS/DE beneficiaries had better outcomes than the non-LIS/DE beneficiaries.

Summary of Examination of the Variation of Within-Contract Differences

- If focusing strictly on a Disability effect, the results are mixed. The majority of the measures examined had a negative within-contract difference and some had no difference or a positive within-contract difference. This means for some measures (negative within-contract difference) the Disabled beneficiaries have worse outcomes as compared to the non-Disabled, but for other measures (no difference or positive difference) the Disabled have similar or better outcomes than the non-Disabled.

Summary of Examination of the Variation of Within-Contract Differences

Median Percentage Point Difference between Groups	LIS/DE - non-LIS/DE	Disabled - non-Disabled	Overlap of Measures
Less than (-2)	8	10	7
Greater than 2	1	3	1
Between (-2) and 2 (inclusive)	7	2	2
Total Measure Count	16	15*	NA

* Readmissions is excluded as it is already adjusted for several factors that could determine Disability status.

- There is a great degree of overlap in the measures that exhibited a negative within-contract difference.
- The measures that demonstrated no disparity for Disabled beneficiaries were a subset of the measures that demonstrated no disparity for the LIS/DE beneficiaries.
- There were more measures that exhibited a positive outcome for Disabled beneficiaries as compared to LIS/DE beneficiaries.

Summary

SUMMARY

The research to-date has provided scientific evidence that there exists an LIS/Dual/Disability effect for a small subset of the Star Ratings measures.

The size of the effect is small in most cases and not consistently negative.

Appendix A: Interpretation Aids

Illustration of Difference in Types of Disparity

Contract	Percentage of Low-SES Enrollees	Overall Average Performance	Low-SES Average Performance	High-SES Average Performance	Within-Contract Difference (Low SES - High SES)
A	17.60%	84.90%	80.00%	86.00%	-6.00%
B	61.50%	51.50%	50.00%	54.00%	-4.00%

Within-Contract Differences:

- *The performance for low- versus high-SES differs by -6.0% for Contract A and -4.0% for Contract B.*
- *The average within-contract difference is -5.0%*

Between-Contract Difference:

- *Contract B has worse performance than A in all categories of performance.*

Illustration of Difference in Types of Disparity

To determine the within-contract disparity, the information for Contract A and B is pooled and the values recalculated.

Beneficiaries Enrolled in Contracts A and B		
Contract	Number of Low SES	Number of Higher SES
A	30	140
B	80	50
Total	110	190

NOTE: From the previous slide, the average within-contract difference (disparity) is **(-5%)**.

Total of Contract A and B	Number of Enrollees	Percentage of Low-SES Enrollees	Overall Average Performance	Low-SES Average Performance	High-SES Average Performance	Total (Overall) SES Difference
Enrollee Information	300	57.90%	70.50%	58.20%	77.60%	(-19.40%)

The goal is to estimate the Within-Contract Disparity.

Average Within-Contract Disparity = Total Difference - Between-Contract Disparity

$$(-5\%) = (-19.40\%) - (-14.4\%)$$

Note: Studies that focus on Total Disparity likely overestimate the magnitude of the effect.

Interpreting Odds Ratios for LIS/DE

An Odds Ratio less than 1 (negative) with asterisks (also denoted in orange font) indicates LIS/DE beneficiaries are statistically significantly less likely to have the recommended care or outcome. **Example:** LIS/DE beneficiaries are significantly less likely to have ART than non-LIS/DE beneficiaries.

An Odds Ratio greater than 1 (positive) with asterisks (also denoted in blue font) indicates Disabled beneficiaries are statistically significantly more likely to have the recommended care or outcome. **Example:** Disabled beneficiaries are significantly more likely to have ART than non-Disabled beneficiaries.

HEDIS Measure	LIS/DE Adjustment Odds Ratio	Disability Adjustment Odds Ratio
Rheumatoid Arthritis Management (ART)	0.85***	1.17***
Controlling High Blood Pressure (CBP)	0.99	1.02

An Odds Ratio without asterisks (or black font) indicates a non-significant association. **Example:** For Controlling High Blood pressure there is no significant association in the outcome for LIS/Dual enrollees as compared to non-LIS/DE enrollees.

NOTE: The number of asterisks indicates the level of statistical significance of the association between LIS/DE and/or Disability and the measure examined.

* Significant at $p < 0.05$

** Significant at $p < 0.01$

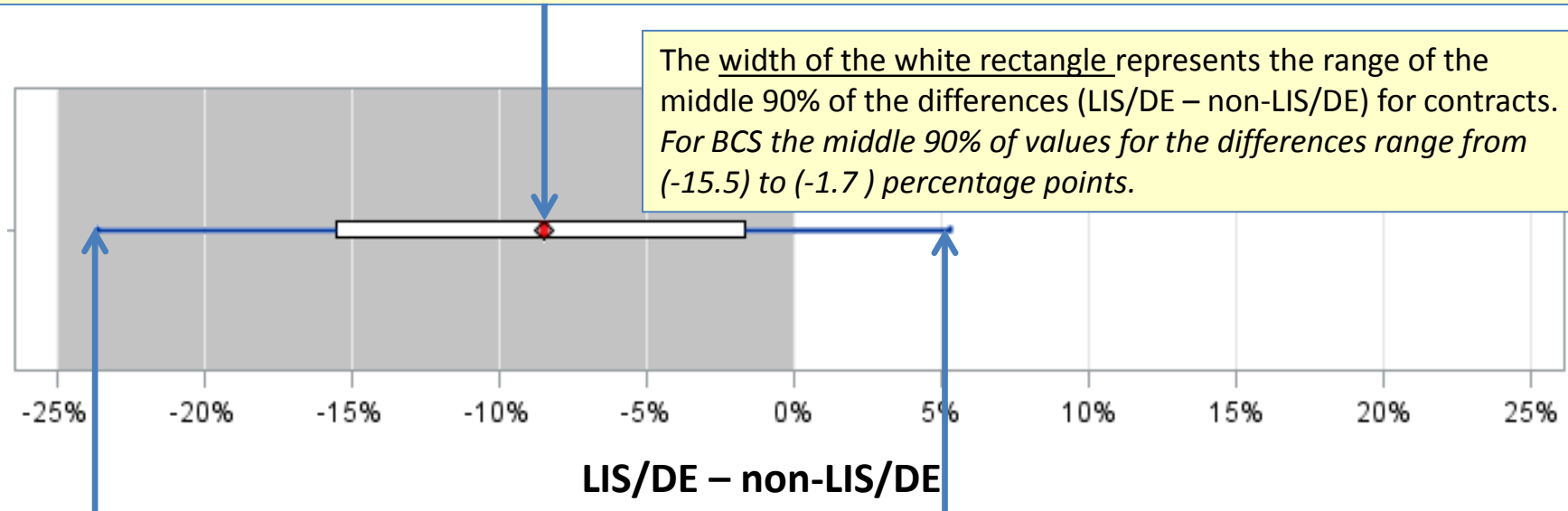
*** Significant at $p < 0.001$

Interpreting the Visual for the Variation in the LIS/DE Disparity Across Contracts

Example: Breast Cancer Screening (BCS)

The red square represents the median difference between LIS/DE and non-LIS/DE rates for the measure for contracts. *For BCS the median difference is (-8.5) percentage points.*

The width of the white rectangle represents the range of the middle 90% of the differences (LIS/DE – non-LIS/DE) for contracts. *For BCS the middle 90% of values for the differences range from (-15.5) to (-1.7) percentage points.*



The estimated minimum value for the difference between LIS and non-LIS for a contract is represented by the first point on the blue line (e.g., for BCS, a -23.6 percentage points difference).

The estimated maximum value for the difference between LIS /DE and non-LIS/DE for a contract is represented by the end point on the blue line (e.g., for BCS, a +5.3 percentage points difference).

The horizontal axis represents the difference in the LIS/DE and non-LIS/DE performance rates per contract.

The estimated (variance adjusted) difference is used to remove the effect of noise (sampling error). Unadjusted differences (without the correction for noise) upwardly bias the range and would result in a larger than true range for the LIS/DE – non-LIS/DE differences.

Appendix B: Additional Research Findings

Associations with LIS/DE Were Smaller for Some Measures When Also Controlling for **Disability**

HEDIS Measure (MA Contracts)	Odds Ratio for LIS/DE with the following in the model:	
	LIS/DE only	LIS/DE and Disability
Adult BMI Assessment	1.11***	1.12***
Rheumatoid Arthritis Management	0.85***	0.82***
Breast Cancer Screening	0.69***	0.75***
Controlling High Blood Pressure	0.99	0.99
Diabetes Care – Blood Sugar Controlled	0.68***	0.75***
Diabetes Care – Eye Exam	0.93***	1.01
Diabetes Care – Kidney Disease Monitoring	0.93***	1.01
Colorectal Cancer Screening	0.87***	0.89***
Osteoporosis Management in Women who had a Fracture	0.71***	0.71***
Plan All-Cause Readmissions [#]	0.87***	N/A
Annual Flu Vaccine	0.85***	0.90***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

[#] Measure is reverse coded to make interpretation of Odds Ratio the same as other measures.

* Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being LIS/DE.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being LIS/DE.

Black Odds Ratio indicates no significant effect. (No * indicator.)

N/A: Not further adjusted for Disability. Readmissions is adjusted for factors that might be part of a person's reason for Disability.

Associations with LIS/DE Were Smaller for Some Measures When Also Controlling for **Disability**

HOS and PDE Measure (MA Contracts)	Odds Ratio for LIS/DE with the following in the model:	
	LIS/DE only	LIS/DE and Disability
Monitoring Physical Activity	0.98	0.98
Reducing the Risk of Falling	1.67***	1.67***
Taking Diabetes Medication as Directed^	0.94***	0.99**
Taking Hypertension Medication as Directed^	0.86***	0.90***
Taking Cholesterol Medication as Directed^	0.94***	0.98***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

Measure is reverse coded to make interpretation of Odds Ratio the same as other measures.

* Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being LIS/DE.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being LIS/DE.

Black Odds Ratio indicates no significant effect. (No * indicator.)

^The sample sizes for PDE were very large, so very small differences become statistically significant.

Associations with Disability Were Smaller for Some Measures When Also Controlling for LIS/DE

HEDIS Measure (MA Contracts)	Odds Ratio for Disabled with the following in the model:	
	Disabled	Disabled and LIS/DE
Adult BMI Assessment	0.93***	0.92***
Rheumatoid Arthritis Management	1.17***	1.20***
Breast Cancer Screening	0.72***	0.75***
Controlling High Blood Pressure	1.02	1.02
Diabetes Care – Blood Sugar Controlled	0.63***	0.65***
Diabetes Care – Eye Exam	0.68***	0.68***
Diabetes Care – Kidney Disease Monitoring	0.69***	0.69***
Colorectal Cancer Screening	0.47***	0.87***
Osteoporosis Management in Women who had a Fracture	0.56***	0.56***
Plan All-Cause Readmissions [#]	1.06***	N/A
Annual Flu Vaccine	0.72***	0.73***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

[#] Measure is reverse coded to make interpretation of Odds Ratio the same as other measures.

* Significant at p<0.05 ** Significant at p<0.01 *** Significant at p<0.001

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being Disabled.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being Disabled.

Black Odds Ratio indicates no significant effect. (No * indicator.)

N/A: Not further adjusted for Disability. Readmissions is adjusted for factors that might be part of a person's reason for Disability.

Associations with Disability Not Affected by Also Controlling for LIS/DE

HOS and PDE Measure (MA Contracts)	Odds Ratio for Disabled with the following in the model:	
	Disabled	Disabled and LIS/DE
Monitoring Physical Activity	1.34***	1.34***
Reducing the Risk of Falling	1.32***	1.33***
Medication Adherence for Diabetes Medications^	0.75***	0.75***
Medication Adherence for Hypertension^	0.72***	0.74***
Medication Adherence for Cholesterol^	0.79***	0.79***

NOTE: Separate analyses conducted for LIS/DE and Disability adjustment. Models control for between-contract differences.

Measure is reverse coded to make interpretation of Odds Ratio the same as other measures.

* Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$

Blue Odds Ratio greater than 1.0 indicates a significant positive effect of being Disabled.

Orange Odds Ratio less than 1.0 indicates a significant negative effect of being Disabled.

Black Odds Ratio indicates no significant effect. (No * indicator.)

^The sample sizes for PDE were very large, so very small differences become statistically significant.

Questions

PARTCANDDSTARRATINGS@CMS.HHS.GOV