

September 23, 2016

Andy Slavitt
Acting Administrator
Centers for Medicare & Medicaid Services
Department of Health and Human Services
200 Independence Avenue, SW
Washington, DC 20201

Dear Acting Administrator Slavitt:

Earlier this year, we encouraged the Centers for Medicare & Medicaid Services to refine the body of measures designed to provide health care consumers with meaningful hospital quality data. At the same time, we questioned the premise of whether the myriad dimensions of reported hospital quality could effectively be reduced into a simple five-point scale.

We were not alone in our concern. Before the July release of hospitals' star ratings, a majority in Congress and many others voiced concerns about the equity of the methodology.¹ At that time, evidence suggested that the rating system could adversely and disproportionately impact safety-net hospitals and, counterintuitively, disadvantage large hospitals electing to report additional measures.

Release of the hospital star-rating system allowed a much more comprehensive analysis of the data. Regrettably, we found that the shortcomings of the methodology — specifically, its lack of recognition of the relationship between social determinants of health and adverse health outcomes — matched our predictions. We are certain that the goal of CMS officials, and other U.S. Department of Health & Human Services stakeholders, is not to disadvantage safety-net providers or hospitals that have robust reporting programs. In fact, CMS and HHS officials are on record on the matter.

Cara James, Ph.D., Director of CMS' Office of Minority Health, recently stated that as much as 80 percent of health disparities are driven by social determinants of health, and that structural barriers are in place to prevent the health care system from effectively addressing these conditions.² Karen DeSalvo, M.D., Acting Assistant Secretary for Health at HHS, also has stated that "Your ZIP code is more important to your health than your genetic code."³ Moreover, a January 2016 report commissioned by OMH, "Guide to Preventing Readmissions Among Racially and Ethnically Diverse Medicare Beneficiaries," acknowledges higher readmission rates for socially complex patients that aren't explained by clinical differences.⁴ This suggests that two hospitals of equal quality, but unequal sociodemographic-status mix, will experience different penalties under the Hospital Readmission Reduction Program.

Further, CMS has signaled its understanding of the influences of social determinants. For example, the Medicare Advantage star-rating incentive program will be adjusted for differences in dual eligibility and disability status in fiscal year 2017.⁵

Given the very real financial and reputational considerations, we analyzed the data supplied by CMS and Yale-CORE. Our goal was to evaluate variation in the star ratings by hospital characteristics, including the socioeconomic status of hospitals' ZIP codes, and to characterize the impact of certain assumptions and components of the measure set. Our research focused on two significant considerations — whether safety-net hospitals' ratings would be influenced by SDS adjustment, and whether case complexity and robust surveillance could bias ratings.

Our analysis yielded clear evidence of a systematic relationship between the number of stars awarded and SDS factors at both the hospital ZIP code and case-mix levels, and between the number of measures reported and domains used in the models. We also uncovered evidence of the model's extreme sensitivity to measures with questionable validity⁶ and measures presenting redundant constructs within the readmissions domain.⁷ Detailed results are included in the attachment, but a high-level review reveals the following.

- A 277 percent difference in a standard socioeconomic deprivation index for the home ZIP codes of one-star and five-star hospitals nationally
- Differences in race, poverty and educational attainment are significant in the home ZIP codes of one-star and five-star hospitals.
- Compared to five-star hospitals, one-star providers had significant differences in supplemental security income ratio, disproportionate share hospital percentage and uncompensated care.
- On average, one-star hospitals reported 50.4 measures in 6.6 domains. This is 39.3 and 15.2 percent points higher than five-star hospitals, respectively.

Given the challenges presented in these findings, we again urge CMS to work with stakeholders to further evaluate and refine the star-rating system. The lack of a unified voice or equitable policy from CMS for the hospital star-rating program is of concern given that existing policies in the hospital quality measures that make up the star ratings disproportionately impact safety-net hospitals. The continued promulgation of quality measures that adversely impact hospitals serving indigent communities is a practice CMS should reverse.

Sincerely,



Herb B. Kuhn
President and CEO

hbk/tls

attachment

c Kate Goodrich, M.D., CMS
Patrick Conway, M.D., CMS
Cara James, Ph.D., CMS
Karen DeSalvo, M.D., HHS

¹ Letter to Acting Administrator Andy Slavitt from the Missouri Hospital Association (2016). Retrieved from <http://www.mhanet.com/mhaimages/SlavittLetter.pdf>

² Modern Healthcare. (2016, April 23). *Q&A: Building the business case for achieving health equity*. Retrieved from <http://www.modernhealthcare.com/article/20160423/PODCAST/304239941>

³ Krisberg, K. (2016). *Social determinants take center stage in call for Public Health 3.0*. *The Nation's Health*, 46(5), 20. Retrieved from <http://thenationshealth.aphapublications.org/content/46/5/20.full>

References

⁴ Betancourt, J., Tan-McGrory, A. & Kenst, K. (2015, September). *Guide to preventing readmissions among racially and ethnically diverse Medicare beneficiaries*. Prepared by the Disparities Solutions Center, Mongan Institute for Health Policy at Massachusetts General Hospital. Baltimore, MD: Centers for Medicare & Medicaid Services Office of Minority Health.

⁵ U.S. Centers for Medicare & Medicaid Services. (2016, April 4). 2017 rate announcement and call letter. Retrieved from <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Announcement2017.pdf>

⁶ Rajaram R., Barnard, C. & Bilimoria, K. (2015). *Concerns about using the patient safety indicator-90 composite in pay-for-performance programs*. *JAMA*; 313(9), 897-898. doi:10.1001/jama.2015.52. Retrieved from <http://jama.jamanetwork.com/article.aspx?articleid=2109967>

⁷ Vella, F. (2016) Comment Letter Prepared for the American Hospital Association. Retrieved from <http://www.aha.org/content/16/16georgetownmeas.pdf>

Attachment: MHA Analysis of the CMS Overall Hospital Quality Star Rating Methodology.

Background: We conducted a series of analyses to explore the hypotheses voiced by numerous concerned stakeholders prior to the July 2016 data release. Using the data and SAS packages supplied by CMS and Yale-CORE, we sought to evaluate variation in the star ratings by hospital characteristics, including the sociodemographic status of the ZIP codes in which hospitals are located. We also sought to test the sensitivity of the star rating models to key measures, assumptions and inclusion criteria used by the measure developers.

Bivariate Analysis: The two most common critiques of the star ratings prior to their release were that hospitals that serve less-advantaged communities would be disproportionately ranked unfavorably because the underlying outcome measures are not adjusted for SDS^{i, ii} and similarly that larger hospitals would fare less favorably because they treat the most complex cases, have more robust surveillance systems and report on more of the measures used by the star ratings.ⁱⁱⁱ

Our analysis yielded clear evidence of a systematic relationship between the number of stars awarded and SDS factors at both the hospital ZIP code and case mix levels (table 1). Many of the area-level SDS factors we evaluated share a monotonic relationship with the number of stars awarded in the direction previously hypothesized by opponents of the overall quality rating system. For example, there is a 277 percent difference in a standard socioeconomic deprivation index for the home ZIP codes of one- and five-star hospitals nationally. One-star hospitals, on average, reside in ZIP codes that are 46 percent non-white and where 17.6 percent of the adult population hold less than a high school education, compared to five-star hospitals with home ZIP code populations that are 24 percent non-white where 9 percent of adults have less than a high school education (a difference of 91 and 92 percent, respectively).

Compared to five-star hospitals, one-star providers had nearly four-fold differences in both Supplemental Security Insurance ratio and Disproportionate Share Hospital percentage. Another indicator of the social and economic contextual surroundings of hospitals' patients is the average amount of uncompensated care per claim—a signal of un- and underinsured payer mix. One-star hospitals faced an average \$3,801 in uncompensated care per claim compared to just \$170 for five star hospitals — a 22-fold difference.

The concern over systematic bias for larger hospitals also garners empirical support from the data. While it is not clear to us why a cut point of three out of seven domains with some representation by outcomes measures was used as the reporting threshold, we question why measure values from hospitals below this threshold were retained in the derivation of the ratings. Regardless of the rationale, a significant relationship appears to exist between the number of stars awarded and both the number of reported measures (of 64) and number represented domains (of 7) used in the latent variable models. This is potentially an artifact of the weighted likelihood approach which increases factor loadings for measures with larger denominators. On average, one-star hospitals reported 50.4 measures in 6.6 domains. This is 39.3 and 15.2 percent higher than five-star hospitals, respectively (table 2). Further, measures of volume, urbanity and case complexity each share a near-monotonic inverse relationship with the number of stars awarded.

Sensitivity Analysis: We analyzed the sensitivity of the latent variable models to the arbitrary inclusion or exclusion of measures by limiting the data to hospitals with all seven domains represented; sensitivity to the exclusion of particular measures with questionable methods^{iv}, and/or questionable unidimensionality^v (PSI-90 and HWR, respectively); and sensitivity to domain completeness with PSI-90 and HWR excluded. We compared the results with the base CMS model in terms of the prevalence of changed star designations for hospitals and calculated measures of interrater reliability.

Compared to the base CMS model, the complete domain model added one star to nearly one-third (871, 29.1%) of all hospitals with seven domains represented, while surprisingly deducting stars from none. The complete domain model also featured modest agreement with the base CMS model (Kappa = 0.57), suggesting the models are very sensitive to the arbitrary inclusion or exclusion of domains and underlying measures (table 3).

Our findings also show that the star ratings are extremely sensitive to the exclusion of methodologically questionable and potentially repetitive individual measures. Excluding PSI-90 from the safety of care domain changed the star designations for 1,350 hospitals (29.7%), with the majority having a star taken away. Removing the single PSI-90 measure yielded results with limited agreement with the base CMS model (Kappa = 0.52). The models were less sensitive to the exclusion of HWR individually, with 15.1 percent of hospitals changing star designations and moderate agreement with the original ratings (Kappa = 0.75).

The final sensitivity test was limited to hospitals with seven domains, and excluded both the PSI-90 and HWR measures. Imposing these assumptions changed the designation of 36.8 percent of included hospitals with a range of two stars lost to three stars gained. This approach also revealed very limited agreement with the base CMS model (Kappa = 0.45), suggesting strong sensitivity of the existing measures to the modifications we tested and raising questions on the reliability of the measures. Additional analysis is needed to identify hospital characteristics associated with positive and negative impacts from the results of these sensitivity tests.

		1 Star	2 Star	3 Star	4 Star	5 Star	1 star vs. 5 star % Difference
Home ZIP Code-Level	Unemployment Rate ¹	12.68	11.30	9.32	8.54	7.47	69.7%
	Percent Non-white ¹	46.04	32.56	22.33	18.81	24.06	91.3%
	Percent Age 25+ Less Than High School ¹	17.58	16.50	14.97	12.01	9.16	91.9%
	Poverty Rate ¹	18.14	15.77	13.19	11.10	10.23	77.4%
	Childhood Poverty Rate ¹	27.51	24.48	21.17	18.14	16.39	67.8%
	Median Household Income ¹	\$47,248	\$46,982	\$48,367	\$53,911	\$58,501	-19.2%
	Residential Vacancy Rate ¹	12.25	11.87	13.48	12.15	9.90	23.8%
SED Index ¹	0.46	0.31	0.13	-0.10	-0.26	276.9%	
Patient Mix-Level	SSI Ratio ²	1,349	952	566	488	361	274.2%
	DSH Percent ³	0.48	0.36	0.30	0.23	0.13	279.6%
	Uncompensated Care per Claim ³	3,801	1,028	867	577	170	2132.7%

Sources: Hospital Industry Data Institute calculations of CMS star ratings data for all hospitals used in the star calculations (i.e. not just those meeting reporting criteria) merged with 2015 Nielsen PopFacts Premier data (1), the 2016 CMS INPpsf1601 file (2) and FY2017 CMS Final Rule Impact File Data (3). Note: The hospital ZIP code socioeconomic deprivation (SED) index is the mean Z-score of the Nielsen variables presented in the table with home value and household income scaled by -1 to maintain additivity.

		1 Star	2 Star	3 Star	4 Star	5 Star	1 star vs. 5 star % Difference
Data Availability	# Domains (of 7) ¹	6.6	6.6	5.7	6.4	5.8	15.2%
	# Measures (of 64) ¹	50.4	47.9	33.7	42.7	36.2	39.3%
Volume and Severity	Beds ²	310	247	170	176	131	136.5%
	Average Census ²	204	151	93	99	69	193.9%
	Transfer-Adjusted Cases ²	3,468	3,429	2,422	2,830	2,324	49.3%
	Outlier Payment % Total Operating Payments ²	6.2%	4.3%	4.4%	4.1%	2.6%	134.8%
	% Large Urban ²	63.1%	42.1%	22.9%	32.0%	33.1%	90.9%

Sources: Hospital Industry Data Institute calculations of CMS star ratings data for all hospitals used in the star calculations (i.e. not just those meeting reporting criteria) (1), and for IPPS hospitals merged with the FY2017 CMS Final Rule Impact File Data (2).

Table 3: Summary Findings of CMS Overall Hospital Star Rating Sensitivity Analysis

		Base CMS Model		Complete Domain Model		PSI-90 Exclusion Model		HWR Exclusion Model		Complete Domain, PSI-90 and HWR Exclusion Model	
		n	%	n	%	n	%	n	%	n	%
		Distribution of Ratings	1 star	141	3.1%	75	2.5%	203	4.5%	105	2.3%
2 star	769		16.9%	445	14.9%	1,074	23.6%	675	15.0%	666	22.2%
3 star	2,505		55.1%	1,196	39.9%	2,522	55.5%	2,520	55.9%	1,340	44.8%
4 star	1,014		22.3%	1,128	37.7%	677	14.9%	1,093	24.2%	763	25.5%
5 star	121		2.7%	150	5.0%	72	1.6%	117	2.6%	104	3.5%
Total	4,550		100.0%	2,994	100.0%	4,548	100.0%	4,510	100.0%	2,994	100.0%
Distribution of Stars Gained or Lost	Lost 2 stars	-	-	0	0.0%	22	0.5%	2	0.0%	12	0.4%
	Lost 1 star	-	-	0	0.0%	1,076	23.7%	221	4.9%	501	16.7%
	No Change	-	-	2,123	70.9%	3,198	70.3%	3,830	84.9%	1,892	63.2%
	Gained 1 star	-	-	871	29.1%	248	5.5%	454	10.1%	553	18.5%
	Gained 2 stars	-	-	0	0.0%	4	0.1%	3	0.1%	35	1.2%
	Gained 3 stars	-	-	0	0.0%	0	0.0%	0	0.0%	1	0.03%
	Total Movement	-	-	871	29.1%	1,350	29.7%	680	15.1%	1,102	36.8%
Kappa		-		0.5743		0.5212		0.754		0.4548	
Weighted Kappa		-		0.6913		0.6288		0.8087		0.5828	

Source: Hospital Industry Data Institute calculations of CMS star ratings data for all hospitals used in the star calculations (i.e. not just those meeting reporting criteria)

References:

ⁱ Letter to Acting Administrator Andy Slavitt from Congress (2016). Retrieved from <http://www.aha.org/advocacy-issues/letter/2016/160331-stardearcolleague.pdf>

ⁱⁱ Letter to Acting Administrator Andy Slavitt from the Missouri Hospital Association (2016). Retrieved from <http://www.mhanet.com/mhaimages/SlavittLetter.pdf>

ⁱⁱⁱ American Association of Medical Colleges (2016). Star Ratings “Deeply Flawed,” Offer Incomplete Picture of Performance. Retrieved from https://www.aamc.org/newsroom/newsreleases/464306/20160727_starratingsrelease.html

^{iv} Rajaram R., Barnard, C. Bilimoria, K. (2015) *Concerns about using the patient safety indicator-90 composite in pay-for-performance programs*. JAMA; 313(9):897-898. doi:10.1001/jama.2015.52. Retrieved from <http://jama.jamanetwork.com/article.aspx?articleid=2109967>

^v Vella, F. (2016) Comment Letter Prepared for the American Hospital Association. Retrieved from <http://www.aha.org/content/16/16georgetownmeas.pdf>