799 9th Street NW
Suite 210
Washington, DC 20001

T (202) 354-2600
vizientinc.com

vizient®

March 29, 2019

Submitted via E-Mail: cmsstarratings@yale.edu

Dr. Kate Goodrich
Chief Medical Officer
Centers for Medicare & Medicaid Services
7500 Security Blvd
Baltimore, MD 21244

Re: Overall Hospital Quality Star Rating on *Hospital Compare* Public Input Request

Dear Dr. Goodrich & Hospital Compare Overall Rating Team:

Vizient, Inc. appreciates the opportunity to respond to the request for public comment from the Centers for Medicare and Medicaid Services (CMS) to gain feedback from stakeholders on several potential updates to, and future considerations for, the methodology of the Overall Hospital Quality Star Rating on *Hospital Compare*. We respectfully submit our comments regarding the specific topics that address changes in hospitals' Overall Hospital Quality Star Ratings. Vizient is pleased to provide input on the agency's plans for longer-term, potential future directions for the Overall Hospital Quality Star Ratings.

Background

Vizient is the nation's largest health care performance improvement company. Our mission is to strengthen our members' delivery of high-value care by aligning cost, quality and market performance. Vizient is member-driven and member-minded, working tirelessly to amplify each organization's impact by optimizing every interaction along the continuum of care. We serve a diverse membership including academic medical centers, pediatric facilities, community hospitals, integrated health delivery networks and non-acute health care providers. Vizient is headquartered in Irving, TX with locations in Chicago, Washington, D.C., and other cities across the country.

Recommendations

People rely on statistical modeling to provide objective assessments about data and to guarantee a level of certainty that the results are simply not due to random chance. To ensure this statistical objectivity is upheld, researchers, data scientists and statisticians must evaluate whether the data and the results meet the necessary modeling requirements; otherwise, like looking in a funhouse mirror, the results become distorted, unstable and less dependable.

Since 2005, Vizient has been using patient data, statistical modeling and outcomes analysis to bring reliable and actionable insights to our member hospitals and their clinicians to help them understand their performance and identify areas where improvement is necessary. Our annual Quality and Accountability Ranking[1] measures performance based on the Institute of Medicine's

[1] The Vizient Quality and Accountability Survey utilizes the following data sources to rate the performance of participating hospitals: the Center for Medicare and Medicaid Services (CMS) Core Measures, the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey, and the Centers for Disease Control and Prevention's National Healthcare Safety Network (NHSN) and the Vizient Clinical Data Base/Resource Manager™

(IOM's) six domains of care: safety, timeliness, effectiveness, efficiency, equity and patient centeredness. Vizient utilizes a composite scoring system for ranking, which uses current, patient-level performance data from a variety of public sources, including the CMS Core Quality Measures and the Vizient Clinical Data Base[1]. In 2018, nearly 400 hospitals participated in the study.

Given Vizient's experience and expertise in analyzing data and rating hospitals in performance measures, the introduction of CMS' Overall Hospital Quality Star Ratings in 2016 was welcomed as another mechanism to help drive performance improvement, while also serving as a resource for patients. Since their introduction, Vizient has been analyzing each update to determine if the methodology used by CMS is meeting the goal of statistical objectivity. Based on the results of our assessments, Vizient has continued to express our concerns[2] that the current methodology is providing unstable results, and has shared these findings and recommendations with the agency.

Vizient continues to urge CMS to discontinue using Latent Variable Modeling (LVM) and instead leverage a more explicit, easier to understand measure weighting – similar to the current precedence the agency has set to use explicit measure weighting in its pay-for-performance programs which Vizient also leverages in our own Quality and Accountability Hospital Ranking methodology. **Furthermore, Vizient urges CMS to remove the publication of the Star Ratings from the *Hospital Compare* website until the agency addresses significant concerns with the methodology.** In doing so, we hope you will consider the recommendations detailed below as well as other, expert feedback regarding the current methodology.

In our comments, Vizient provides specific responses to various issues raised regarding the proposed enhancements to the Overall Hospital Quality Star Rating, as well as the current methodology used in calculating hospital scores. We offer our feedback and recommendations to constructively improve the Overall Star Rating project. Our comments reflect the views of our organization, as well as input received from our hospital members from across the nation.

Assessment and Recommendations

In November 2018, CMS announced two methodology updates; to remove measures with statistically significant negative-loading coefficients and to change the weighting of hospital-associated infection measures in the safety-of-care group. However, CMS' two new methodology updates, which were reported in the February 2019 release, do not address concerns voiced by Vizient, as well as other stakeholders, particularly regarding the latent variable modeling approach. While the intent may have been to address methodology issues, CMS has instead inadvertently potentially introduced even more instability into the Hospital Star Rating system. Vizient is extremely concerned that the flawed methodology currently used to determine the Ratings that are posted on the *Hospital Compare* website are both inaccurate and misleading to patients seeking care.

Latent Variable Modeling Affects Loading Coefficients to Create Misleading Results

CMS has stated that latent variable modeling provides an objective way to assign measured importance or weights for each of the seven performance areas in the ratings. However, after analyzing the February 2019 publically available CMS Hospital Star Rating data (the most current available) which included two methodological improvements, Vizient continued to identify significant opportunities in the CMS latent variable modeling choices indicating modeling selection bias, producing unreliable loading coefficients and ultimately potentially misleading results.

[2] Vizient Public Comment Letters Submitted on May 30, 2018 and September 27, 2017.

Change 1: Removing Measures with Statistically Significant Negative Loading Coefficients
While the July 2018 Star Ratings were never officially released, hospitals with better performance in the Healthcare-Associated Infection (HAI) measure were being penalized. This counter-intuitive measure evaluation was driven by applying latent variable modeling which calculated negative measure weighting, lowering a hospitals' score for better performance. To address this concern, CMS committed to removing statistically significant measures which penalized hospitals for better performance, otherwise known as negative measure loading coefficients.

For the February 2019 Star Rating release, no measures met the statistically significant criteria as shown in Figure 1[3], but one measure OP-32: Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy measure reports a negative-loading coefficient of -0.01 (Table 1). While this measure has marginal impact on the overall hospital score, the presence of this non-significant negative loading is symptomatic of a sub-optimal modeling approach. Vizient strongly supports the movement to value-based care, and does not believe that hospitals should be penalized by any amount – small or large – for providing better care.

Table 1: Readmission Group February 2019 Measure Loading Coefficients (Data published in CMS February Report)

Measure	Loading Coefficient
Excess Days in Acute Care after Hospitalization for Acute Myocardial Infarction	0.338
Coronary Artery Bypass Graft (CABG) 30-Day Readmission Rate	0.3154
Chronic Obstructive Pulmonary Disease (COPD) 30-Day Readmission Rate	0.5515
Excess Days in Acute Care after Hospitalization for Heart Failure	0.4544
Hospital-Level 30-Day All-Cause Risk- Standardized Readmission Rate (RSRR) Following Elective Total Hip Arthroplasty (THA)/Total Knee Arthroplasty (TKA)	0.4107
Excess Days in Acute Care after Hospitalization for Pneumonia (PN)	0.4372
Stroke (STK) 30-Day Readmission Rate	0.5306
HWR Hospital-Wide All-Cause Unplanned Readmission	0.9976
Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy	-0.01311

Change 2: Healthcare-Associated Infection (HAI) measures
CMS' second methodological shift was to use device days, number of procedures, and patient days instead of predicted infections to weight measure scoring for the HAI measures. CMS stated that the denominators help stabilize the measure weighting within the group and reduces the sensitivity of the methodology to an individual measure change[4]. This methodology update was as a result of the significant loading coefficient swings in the Safety group for the Patient Safety Composite Measures (PSI-90) and the Total Hip & Knee Complications (THK) between the July 2018 (not released) and the December 2017 (released) Star Ratings. Vizient found this to be of considerable concern, as no prior release of the CMS Star Ratings have had the significant shifts we saw in July 2018 as shown below in Table 2.

[3] Figure 1. Measure Selection Flowchart (February 2019)
[4] CMS Data Release Frequently Asked Questions - February 2019. 02/28/19.

Table 2: Comparing Safety Group Loading Coefficients Over Time[5]

Measures	Feb 2019	July 2018 (not released)	Dec 2017	Oct 2017	Dec 2016	Oct 2016	Jul 2016	Apr 2016	Apr 2015
PSI-90-Safety	0.90060	0.16520	0.94420	0.920	0.930	0.930	0.780	0.920	0.770
Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA)	0.19730	0.96400	0.21100	0.210	0.170	0.170	0.680	0.190	0.560
Central-Line Associated Bloodstream Infection (CLABSI)	0.00600	0.02421	0.01836	0.030	0.050	0.060	0.100	0.070	0.020
Catheter-Associated Urinary Tract Infection (CAUTI)	0.00744	-0.00400	0.00094	0.010	0.050	0.110	0.090	0.110	0.100
Surgical Site Infection from colon surgery (SSI-colon)	0.04659	-0.04424	0.04946	0.050	0.120	0.090	0.050	0.090	0.100
Surgical Site Infection from abdominal hysterectomy (SSI-abdominal hysterectomy)	0.06504	-0.01476	0.04735	0.020	0.060	0.060	0.070	0.080	0.090
MRSA Bacteremia	0.03680	0.03397	0.07112	0.080	0.020	0.010	0.560	0.030	0.050
Clostridium Difficile (C.difficile)	0.02582	0.02580	0.01161	0.020	0.030	0.001	0.770	0.000	0.070

Understanding the Effect of the February 2019 Changes

The February Star Rating Safety group loading coefficients appear to be more aligned with previous releases. To better understand the impact of using patient days and device days instead of predicted HAIs, Vizient simulated the impact by analyzing the December 2017 Star Rating publically available data from QualityNet and *Hospital Compare*, substituting patient days and device days for predicted infection in a latent variable modeling algorithm. The results indicated nearly zero changes in Hospital Star Ratings – as only one hospital's rating moved from 3 to 4 stars (Table 3).

Table 3: Vizient Simulated December 2017 Star Ratings versus CMS Published December 2017 Star Rating Comparison (Vizient Analysis using December 2017 CMS data)

		Vizient updated December 2017 Analysis					
		Star-1	Star-2	Star-3	Star-4	Star-5	Total
CMS December 2017 Star Rating	Star-1	261	0	0	0	0	261
	Star-2	0	752	0	0	0	752
	Star-3	0	0	1188	1	0	1189
	Star-4	0	0	0	1153	0	1153
	Star-5	0	0	0	0	336	336
	Total	261	752	1188	1154	336	3691

Vizient compared the impact on the Safety group loading coefficients from our simulations with the published December 2017 coefficients, and found marginal differences as shown in Table 4.

[5] https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228775957951

Table 4: Safety Group Measure Loading Coefficient Comparison: Vizient Simulated versus December 2017 Published (CMS December 2017 Star Publically Available Data Set)

Measures	Dec 2017	Updated Dec 2017*	Difference
PSI-90-Safety	0.94420	0.944073	-0.0001
Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA)	0.21100	0.210817	-0.0002
Central-Line Associated Bloodstream Infection (CLABSI)	0.01836	0.022699	0.0043
Catheter-Associated Urinary Tract Infection (CAUTI)	0.00094	0.002411	0.0015
Surgical Site Infection from colon surgery (SSI-colon)	0.04946	0.045498	-0.0040
Surgical Site Infection from abdominal hysterectomy (SSI-abdominal hysterectomy)	0.04735	0.045425	-0.0019
MRSA Bacteremia	0.07112	0.061520	-0.0096
Clostridium Difficile (C.difficile)	0.01161	0.016985	0.0054

The CMS February methodology document does not reference additional methodological changes that account for the sizable shift in measure loading between July 2018 and February 2019 Star Ratings as shown in Table 2. Coupled with the Vizient simulated results indicating marginal Star Rating changes due to the methodology updates, the February results are disconcerting. We believe the dramatic differences found are due to CMS' continued use of latent variable modeling.

Confusing Results Created by the Current 1-Factor Latent Variable Model Approach

Given the confounding results produced by CMS' latent variable modeling approach, Vizient conducted an in-depth statistical assessment to better understand the methodological issues. CMS currently uses what is known as a 1-factor, weighted latent variable modeling approach, which simply assigns a single weight to each measure. More complex approaches exist, such as 1-factor reduced measures – which only includes measures that are statistically significant – or 2-factor modeling – which assigns two measures weights for a single measure. Vizient closely examined four common model fit statistics used in evaluating latent variable modeling performance, and identified model fit performance opportunities across 4 of the 7 measure groups.

One common model fit statistic, the goodness-of-fit test, assesses how well the latent variable model-generated results compare with the observed data. When simulating model performances 100 times and assessing the goodness-of-fit results, Vizient identified model problems with six of the seven measure groups. The root mean square error approximation is another technique for assessing model performance where a small error value is desirable; however, both the patient experience and the process timeliness groups indicate larger than acceptable model error values.

Another model performance evaluation, the comparative fit index, assesses if the model performs consistently with the observed values. For the February 2019 data, six of the seven groups indicated poor performance with less than 0.95 performance. Lastly, the standardized root mean square residual measure evaluates differences between the actual observed data compared to the model's output. Based on this assessment, Vizient found that four of seven measure groups indicate opportunity for improvement. The combined assessment across all measures can be found below in Table 5.

Table 5: February 2019 CMS Latent Variable Modeling Assessment
Vizient model fit statistics were generated using 1-factor, non-weighted latent variable instead of 1-factor, weighted, due to limited model fit analysis provided by CMS

Groups	Goodness of Fit[1]	Root Mean Square Error Approximation [2]	Comparative Fit Index[3]	Standardized Root Mean Square Residual [4]	# of Statistical Assessments Indicating Modeling Opportunity
Mortality	60%	0.0357	0.9658	0.0543	1/4
Readmission	100%	0.0622	0.9093	0.0647	2/4
Safety	64%	0.0372	0.7601	0.0724	3/4
Patient Experience	100%	0.2122	0.8546	0.0502	3/4
Efficiency	17%	0.0367	0.9349	0.0512	0/4
Process Timeliness	100%	0.1544	0.8807	0.1320	4/4
Process Effectiveness	100%	0.0493	0.7624	0.0881	3/4

[1]Goodness of Fit: to minimize sample size bias, 100 random samples of 500 were taken, percent reflects total of samples with <0.05 p-value indicating significant differences between actual model distribution and model performance.
[2]Root Mean Square Error Approximation ≥0.1 may indicate a serious model fit problem
[3]Comparative Fit Index > 0.90, >0.94, >0.95 have been used as cut-offs of "acceptable" model fit. Less than 0.9 indicates potential opportunity.
[4]Standardized Root Mean Square Residual < 0.055 suggest "good" fit, greater than, equal to 0.055 indicates opportunity

To gain insight into how CMS could potentially improve the latent variable modeling approach, Vizient explored alternative, more complex latent variable modeling approaches to improve model performance – including 1-factor-reduced measures, 2-factor and 2-factor-reduced measures modeling approaches. Vizient found through the various modeling approaches that, while model performance improved per the four model fit statistics referenced, the increased model complexity resulted in lower user interpretability. **While these more complex approaches may be more statistically appropriate than the current CMS 1-factor latent variable modeling approach, the additional complexities intrinsically linked would make it even more difficult for the public and providers to understand.**

Based on our assessment, CMS is in a statistical predicament where the current 1-factor modeling approach does not create reliable results, but the more statistically appropriate techniques are too complicated to understand. **To mitigate this analytical conundrum, Vizient encourages CMS to consider a more simplistic approach, similar to existing pay-for-performance programs** (e.g., the Hospital Readmissions Reduction, Value-Based Purchasing, and Hospital-Acquired Condition Programs). In these programs, the measure sets are clearly defined, with standard weights for each measure evaluated. Vizient strongly believes that this approach would not only improve scoring understanding, but also provide consistency among the CMS performance evaluating programs. Vizient's recommendations are intended to improve the Star Rating's accuracy and clarity for patients, as well as to create important feedback for providers for performance improvement.

Pay-for-Performance Measures and Star Ratings Yield Inconsistent Results

CMS sets the nation's standards for health care performance evaluation through their pay-for-performance strategy and programs. The measures included in the Hospital Value-Based

Purchasing Program (VBP), Hospital-Acquired Conditions Reduction Program (HACRP) and Hospital Readmissions Reduction Program (HRRP) all financially penalize hospitals who do not meet CMS-established performance thresholds.

The measures used in these pay-for-performance programs also contribute significantly to the Overall Hospital Quality Star Ratings – in particular the Readmission, Safety, Patient Experience and Mortality group scores. These groups collectively represent 88 percent of the overall score; however, despite the overall measure alignment, the results between the pay-for-performance and the Star Ratings are inconsistent.

For example, for the fiscal year (FY) 2019 VBP measure noted in Table 6, 1,229 hospitals received a financial penalty for performance. Of those, 452 hospitals received an 'Above the National Average' classification for the Safety group, 105 hospitals received 'Above the National Average' for the Mortality group, and 151 hospitals received 'Above the National Average' for the Patient Centeredness group in the Star Ratings program. Further, of the 2,587 hospitals who received a payment penalty in the HRRP, 945 hospitals also received 'Above the National Average' for the Readmission group in their Star Rating.

Table 6: Hospitals Penalized in the CMS Pay-for-Performance Programs versus the February 2019 Hospital Star Rating 'Above National Average' and Overall 5-Star Performance

CMS P4P FY 2019 Program Penalty	Total Hospitals Penalized	National Comparison Group	Above the National Average	Overall 5 Star Rating
VBP	1229	Mortality Safety Patient Experience	105 452 151	6 25 12
HACRP	800	Safety	161	27
HRRP	**2587**	**Readmission**	**945**	**132**

Vizient believes this is due to methodological differences between the two CMS-supported programs. For the HRRP, CMS evaluates hospitals using quintile binning based on the percent of dual-eligible Medicare payers; whereas, for the CMS Star Rating Readmission group score, no adjustment is made. This disconnect in methodology between the two programs is not only financially penalizing providers, but also affecting the reputation hospitals have worked diligently to earn in each of their communities. Furthermore, it adds to public confusion as to which hospitals are providing the best quality care.

At the individual measure level, the methodological inconsistencies also appear. In the December 2017 Hospital Star measure loading coefficients as shown below in Table 7, the latent variable modeling approach deemed HAIs as non-statistically significant loading coefficients – yet important enough to put hospitals at financial risk for poorer performance.

Table 7: December 2017 Measure Loading Coefficients

Parameter	Estimate	Standard Error	Pr > |t|
Central-Line Associated Bloodstream Infection (CLABSI)	0.01836	0.01149	0.11
Catheter-Associated Urinary Tract Infection (CAUTI)	0.00095	0.0107	0.9308
Surgical Site Infection from abdominal hysterectomy (SSI-abdominal hysterectomy)	0.04736	0.02416	0.0501
Clostridium Difficile (C. difficile)	0.01161	0.008501	0.1719

Vizient found similar results as shown for the February 2019 Safety measure loading coefficients with non-significant p-values for central-line associated bloodstream infection (CLABSI) and catheter-associated urinary tract infection (CAUTI). To rectify these issues, Vizient recommends CMS take a consistent hospital evaluation approach by first assessing the precedents CMS has set in existing pay-for-performance programs, and aligning and streamlining them with the Star Rating methodology.

Lack of Hospital Stratification Limits Usefulness of Star Ratings & Creates Hospital Bias

Vizient encourages CMS to adopt a hospital stratification approach. CMS introduced hospital groupings based on the percentage of dual-eligible patients, which evaluates hospital readmission performance relative to hospitals with similar patient challenges (Quintile-1 represents the lowest percent of dual-eligible, Quintile-5 represents the highest). When socio-demographic status (SDS) is not incorporated into the scoring methodology, hospitals with a higher proportion of complex patients have lower hospital Star Ratings. All of our members believe and practice that every patient who seeks care should receive the same high-quality care. We encourage CMS to monitor this issue for potential unintended consequences, and continue to look for ways to adjust for the risk that some hospitals face due to the proportion of vulnerable patients that they serve. As shown below in Table 8, hospitals with the highest percent of dual-eligible (Quintile-5) patients earn 1-star in the CMS Star Rating program indicating that the current CMS Star Rating program lacks appropriate adjustment for not only patient socio-demographic challenges, but also is limited in the current methodology's ability to account for patient clinical severity or complexities.

Table 8: Percent of Hospitals in FY 2019 CMS Readmission Reduction Program Quintiles versus February 2019 Hospital Star Distribution

		CMS HRRP Quintiles				
		1	2	3	4	5
CMS Hospital Star Ratings	1	8%	13%	19%	23%	37%
	2	9%	17%	19%	25%	30%
	3	12%	23%	23%	24%	18%
	4	26%	25%	22%	17%	10%
	5	47%	23%	16%	7%	7%

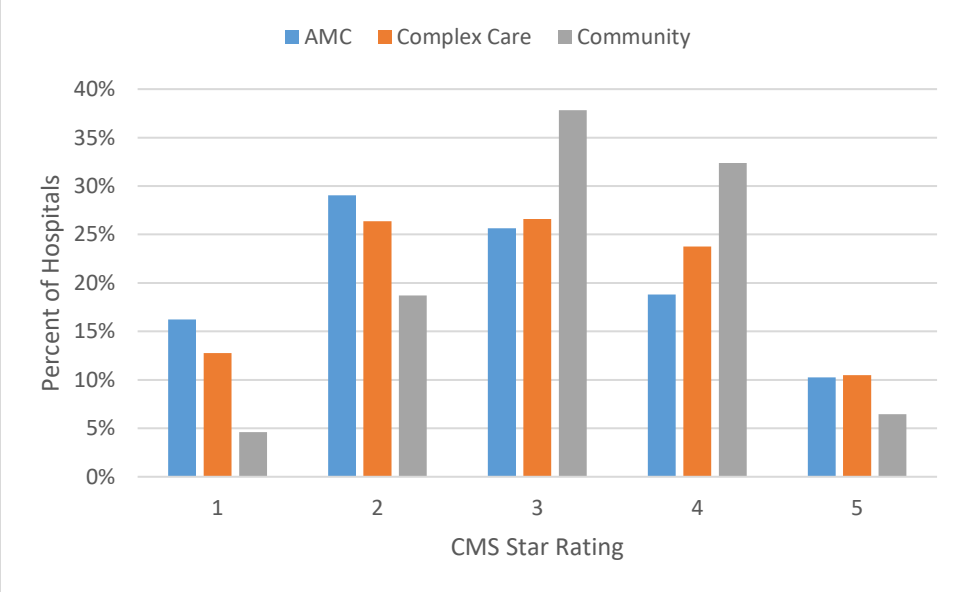
This is especially notable in Academic Medical Centers (AMCs) – with a high proportion represented in the Quintile 4 and 5 percent dual-eligible categories, as compared to Community Hospitals (COMM) and Complex Teaching Medical Centers (CTMC), as shown in Table 9, representing the percentage of hospitals in each cohort that fall into the CMS quintiles used in the HRRP.

Table 9: Percent of Hospitals in the CMS HRRP Quintiles versus Hospital Cohort

	CMS Readmission Reduction Program Quintiles				
	1	2	3	4	5
AMC	10%	20%	22%	27%	22%
CTMC	19%	17%	19%	22%	23%
Comm	23%	25%	22%	16%	14%
Total	20%	20%	20%	20%	20%

As a result, academic medical centers have a higher proportion of 1 and 2 star hospitals in the February 2019 Star Rating distribution chart (graph 1), compared to their community hospital counterparts – indicating an unintentional bias against hospitals with more complex patient populations. Vizient reiterates that safety-net hospitals have other unmeasured differences in patient characteristics that may contribute to differences in readmission rates[6]. SDS factors in risk adjustment allows for fair cross-provider comparisons and does not penalize one provider over another – or give the impression that one provider provides lower-quality care simply due to their ability and readiness to treat any patient. We urge CMS to utilize methodology that encourages equitable care delivery, while also accounting for the disproportionate penalties for safety-net providers and academic medical centers.

Graph 1: February 2019 CMS Hospital Star Rating Distribution by Hospital Cohort



[6] U.S. Department of Health & Human Services, Office of the Assistant Secretary for Planning and Evaluation. "Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs". December 21, 2016.

Vizient identified additional methodological imbalances that resulted by including specialty hospitals in the Overall Hospital Star Ratings. Approximately 44 specialty hospitals were listed, including orthopedic, heart and vascular, and a rehabilitation hospital. Despite their small representation in the overall *Hospital Compare* data, 30 of the 44 (68 percent) received a 5-Star Rating. While it is certainly important to evaluate performance for these hospitals, combining such a heterogeneous mix of hospitals limits the Star Rating's meaningfulness and value for patients.

As Vizient shared in our September 2017 **comment letter to CMS**, until the appropriate hospital cohorts are defined within the CMS Star Rating methodology, hospitals with more complex, tertiary or quaternary care will be unfairly labeled as providing sub-par care. Vizient strongly urges CMS to ensure that safety-net and outlier hospitals are not disproportionately impacted – and recognize that these hospitals treat the most vulnerable and complex patients. Additionally, Vizient detailed the Quality and Accountability framework it utilizes for setting hospital cohort criteria to create meaningful and actionable benchmarks and comparisons for its hospital members. This criteria includes relevant volume thresholds that differentiate patient comorbidities and surgical complexity – including the number of solid organ transplants, cardiac surgery and neurosurgery cases, acute transfers in from other hospitals and trauma service line volume.

Leveraging this criteria, we created a third hospital ranking group to accompany our current Comprehensive Academic Medical Center and Community Hospital groupings. This third cohort, the Complex Care Medical Center group, represents large, complex organizations that focus on special patient services or care, such as safety-net or specific surgical populations. The criteria used to define these cohorts was identified using a combination of exploratory data analysis (measure correlations) of hundreds of data elements and further refined using robust clustering analysis and subject matter expertise to arrive at four main volume driven criteria – solid organ transplants, acute transfers-in, trauma cases, cardiothoracic and neurosurgery volumes. This cascading criteria further supports more meaningful comparisons for hospitals taking care of unique patient needs. Additionally, Vizient developed a separate ranking framework and measures to support critical access hospitals (CAHs), oncology-specific medical centers and pediatric hospitals, which will be introduced in 2020. By splitting hospitals into relevant cohorts, Vizient's modeling more accurately reflects a specific hospital's performance and corresponding rating.

Vizient tested an alternative methodology on the February 2019 Star Ratings data utilizing both clear, standardized weights and appropriate hospital groupings. The standardized weights provide transparency into the rating process and offer a replicable formula hospitals can follow as they work towards tangible improvement. To account for missing or low-volume denominators, Vizient re-allocated the weight from that measure equally to the other measures within that domain. This ensures a fair and balanced score can be achieved for all hospitals.

Additionally, hospitals grouped into cohorts based on the complexity of the patients treated is a key recommendation in order to provide more actionable and reliable hospital comparisons. The three groups used by Vizient were Comprehensive Academic Medical Centers, Complex Teaching Medical Centers, and Community Hospitals referenced above. Critical Access Hospitals and hospitals solely focused on specialty care, such as orthopedics or cardiovascular care, provide a different level of care from the other hospitals in this analysis and were therefore removed. Vizient recommends the development of another group or groups specifically tailored to these unique hospitals. Under the existing Star Rating methodology, hospitals providing more complex, tertiary or quaternary care are artificially labeled as providing below average care, as shown in Graph 1. Splitting the hospitals into relevant groups also provides a weighting adjustment by only comparing hospitals to a population of their peers. By separating hospitals into homogenous cohorts, Vizient is able to offer a more accurate look at a specific hospital's performance and recommends CMS consider a similar approach.

Data Lags & Limited Patient Population Demographic Limitations

Many of the heavily weighted CMS measures, such as the measures in the readmissions, safety and mortality groups are considerably dated, some going as far back as 2014 for the collected measures reported in the February 2019 CMS Hospital Star Ratings. Vizient continues to believe that the current CMS Star Ratings do not reflect current hospital performance, which limits the usefulness of the Star Ratings for patients making health care choices. Furthermore, placing increased weights on these measure groups containing two-year old performance data is misleading to the public by not accurately reflecting the current performance, or as close to current performance data as possible, for measures that are highly visible and of high importance to patients. Vizient encourages CMS to rely on consistent measure and measure group weighting that is updated on an annual basis.

Additionally, because the Star Ratings leverage Medicare data, which represents approximately 10-15 percent of a hospital's total patient population, it primarily focuses on conditions and procedures for the 65 years or older patient populations.

Responses to Specific CMS Questions

The remainder of our comments will address the specific questions that CMS has asked in its Public Input Request[7] on the Overall Hospital Quality Star Rating.

Section 4.1 Measure Grouping

1. *We would like to use a three-step approach (clinical coherence, confirmatory factor analysis, and ongoing monitoring) to define measure groups. Is this approach reasonable?*

Vizient supports adequate measure assessment and groupings based on clinical coherence, preliminary measure analysis and ongoing monitoring measures for clinical relevance and performance opportunity. Additionally, we support CMS' proposal to use confirmatory factor analysis to determine if latent variable modeling is the appropriate statistical approach.

2. *Should CMS use the balance and consistency of loadings as a factor in evaluating grouping?*

Vizient supports, first and foremost, a more clinically grounded approach by leveraging a well-represented clinical expert panel to identify relevant measures and define clinically meaningful groupings. Vizient cautions CMS in using the balance and consistency of the measure loading coefficients as a measure grouping criteria for several reasons. Firstly, selecting measure groupings based on statistical criteria is likely to misalign with clinical groupings which limit grouping relevance and validity. Secondly, from one reporting period to the next, the model may produce inconsistent measure loading results, ultimately introducing additional measure fluctuations and inconsistences to the ratings. Finally, measure loading imbalance may be continue regardless of how measures are grouped. Indicating the modeling approach may not be appropriate for the given data and in turn, CMS would be faced with exploring alternative measure loading approaches which, again, add variability and inconsistency to the ratings.

Section 4.2 Regrouping of Measures

1. *Is the current grouping or one of the potential alternative groupings of the Safety of Care measures most suitable for the Overall Hospital Quality Star Rating based on previously mentioned criteria?*

[7] Overall Hospital Quality Star Rating on *Hospital Compare* Public Input Request. Prepared by: Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE). February 2019.

Vizient recommends continuing with the existing measure groupings, and substituting latent variable modeling for a more explicit, easy to understand measure weighting approach.

Section 4.3 Incorporating Precision of Measures

1. *Do you have any concerns about changing the methodology to use a combination of denominator weighting and log (denominator) weighting, based on the type of measure?*
2. *Do you have any concerns about applying a change to the weighting approach across all measure groups (where data are available) vs. applying the change only to measure groups that meet specific criteria?*
3. *Are there other approaches that CMS should consider?*

Vizient believes measure precision choices limited Star Rating result accuracy given the latent variable modeling challenges we identified. While Vizient acknowledges that measure precision can be improved by incorporating increased denominator weighting or applying various weighting approaches, this change may not be effective in improving the latent variable modeling accuracy or fit. In turn, Vizient recommends exploring a more explicit measure weighting approach and discontinue latent variable model derived measure weighting.

Section 4.4 Period-to-Period Star Ratings Shifts

1. *What are possible benefits and drawbacks to increasing stability by limiting change in this way?*
2. *Should the Overall Hospital Quality Star Rating methodology be modified to incorporate data from previous periods through a time averaged approach?*
3. *Are there other approaches to this CMS should consider?*

Vizient continues to believe the root cause of the period-to-period variation is driven by measure loading coefficients generated from using latent variable modeling – and exacerbated by the current lack of hospital groupings. Trying to smooth out that variation by simply blending the old rating and the new rating is not an effective solution. Many of the heavily weighted CMS measures, such as the measures in the readmissions, safety and mortality groups are considerably dated, some going as far back as 2014, for the collected measures reported in the February 2019 CMS Hospital Star Ratings. Incorporating data from a previous time period would further limit the utility of the Star Ratings. Vizient recommends the use of more timely data, more stable measure weighting approach and creating hospital groupings to minimize period-to-period Star Rating shifts – which would offer a more contemporary look at how a hospital is currently performing.

Section 4.5 Peer Grouping

1. *Would it be valuable to calculate Overall Hospital Quality Star Ratings among peer groups? How should the information be displayed? If CMS decides to move forward with this feature, which stakeholders do you believe would use the information and how would they use it?*

Given the methodological limitations and imbalanced evaluation of various hospital types within the same scoring framework, Vizient fully supports hospital or peer groupings and has outlined our recommendations above regarding a robust process for identifying differentiating hospital characteristics based on patient acuity and complexities – as well as the depth and breadth of services offered.

Vizient suggests a simplistic approach to displaying hospital star ratings. While CMS could certainly explore displaying a ‘Top Hospital’ within each hospital peer group, this effort may be unnecessary as long as the public is aware the hospital is recognized as a ‘Top Hospital’. Additional acknowledgements or creation of a second ‘Overall’ star rating would be unnecessary. CMS has experienced success with hospital grouping using the percentage of

dual-eligible patients in the HRRP, and Vizient encourages CMS to explore similar approaches for the Star Ratings.

As shown in our CMS Hospital Grouping assessment, hospitals grouped with like-hospitals (i.e., their peers that offer similar services and care for similar patients) are evaluated in a more consistent, robust and comparable way that provides clearer insight into performance for both providers and the public.

2. *Among the feasible variables that could be used for peer grouping (specialty, number of measures reported, teaching status, number of beds, critical access hospital, proportion of dual eligible patients), which would be most useful? Descriptions for each mentioned variable are included below.*

    a. Proportion of dual-eligible describes the proportion of patients eligible for both Medicare and Medicaid. Dual-eligible could be used to peer group hospitals with similar proportions of dual-eligible patients by quintile, for example.
    b. Teaching hospitals are those that have one or more accredited residency programs or have an intern or resident to bed ratio of 0.25 or higher. Teaching and non-teaching hospitals may differ in mission, financial considerations, and services. Teaching status could be used to peer group teaching and non-teaching hospitals.
    c. Number of beds at a hospital is a proxy for hospital size. Smaller hospitals may have fewer services and resources while larger hospitals tend to be in urban areas and may serve disadvantaged populations.
    d. Hospitals that report more measures may not be directly comparable to hospitals that report fewer measures. Number of measures reported could be used to group hospitals by quartile, for example.
    e. Certain rural hospitals can qualify as critical access designation for CMS purposes to indicate lack of proximity to other hospitals for prospective patients. Hospitals could be grouped as either critical access or non-critical access.
    f. Specialty hospitals are those that that primarily or exclusively engage in the care and treatment of patients with cardiac conditions, orthopedic conditions, conditions requiring surgical procedures, or other specialized services. Hospitals could be grouped and compared as specialty or non-specialty.

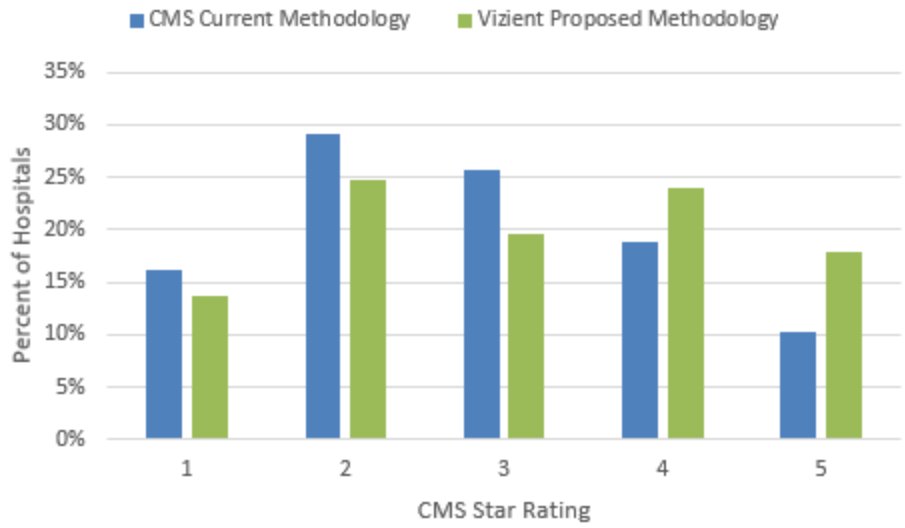
While we appreciate CMS' possible peer grouping scenarios, we strongly believe that hospital peer grouping should be based on relevant volume thresholds that differentiate patient comorbidities and surgical complexity: the number of solid organ transplants, cardiac surgery and neurosurgery cases, acute transfers in from other hospitals and trauma service line volume. In arriving at these criteria, Vizient explored many of the options provided and found that no one single factor or characteristic provides sufficient separations or adjustments for hospitals differences. The proportion of dual-eligible patients, while insightful and relevant for readmissions and excess days measures, does not fully represent the severity or complexity of patients as would transfer in status or trauma case volume. Similarly, for teaching status or number of beds, these characteristics provide some insight, but given the variety of teaching programs and the different severity of the types of patients, Vizient found these criteria, used in isolation, were limited in creating 'like-hospitals'.

CMS' recommendation to evaluate measures reported is a step forward toward evaluating the types and volume of patients seen by the hospitals, but would not necessarily adjust the differences across measures reported. For instance, if hospital A reports three heart failure measures and hospital B reports three surgical complication measures, the comparison in outcomes may not be as relevant.

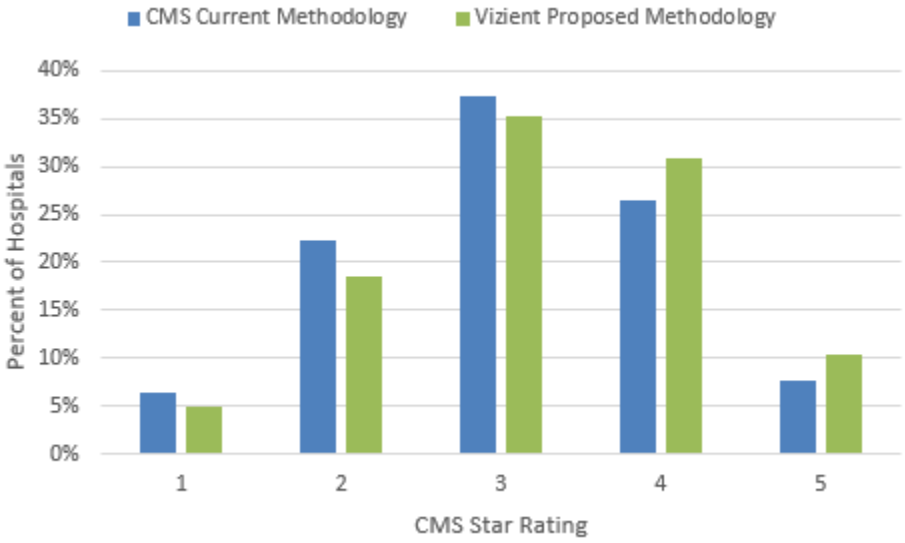
Vizient applied our recommended approach of grouping hospitals to the CMS February 2019 data. Additionally, we removed critical access and specialty hospitals from the assessment and

weighted the measures equally. In comparing the February 2019 CMS to the Vizient Hospital Groupings, AMCs are more evenly represented in the 4 and 5 Star Ratings (graph 2), and Complex Teaching Medical Centers and Community hospitals have only a slight adjustment in Star Ratings (graph 3). Vizient believes that this approach provides a more practical, comparable assessment of hospital performance that limits bias due to limited measure representation or differences in full hospital patient acuity.

Graph 2: AMC February 2019 Star Rating Distribution Comparison with Vizient Proposed Methodology



Graph 3: Community February 2019 Star Rating Distribution Comparison with Vizient Proposed Methodology



Section 5.2 Explicit Approach

1. *What are the advantages and disadvantages of a more explicit approach to calculating Overall Hospital Quality Star Ratings?*
2. *Is the explicit approach a worthwhile change in approach and direction to consider further?*
3. *How could such an approach be best operationalized or sustained?*

Despite the various adjustments and alternative ideas to improve the latent variable modeling approach, the measure loading coefficients continue to generate inconsistent and clinically counter-initiative results by penalizing hospitals that provide better care. Vizient is supportive of CMS exploring alternative measure weight approaches and supports an explicit measure weighting approach due to its clear, straight-forward application, which will be easy for providers and the public to understand. Grouping hospitals by complexity of patients seen and services provided creates inherent weighting adjustments by simply comparing a hospital to members of its peer group. When a hospital is missing a measure, Vizient suggests distributing the weight from the missing measure to the other measures within the domain. A minimum number of measures would be required to receive a score in that domain. Vizient believes any disadvantages, such as measures being removed or differences in hospital volume, in the explicit weighting approach pale in comparison to the challenges CMS has faced by using latent variable modeling. **Thus, coupled with hospital peer grouping, Vizient recommends CMS explore an equal weighting approach that is similar to those used in the current pay-for-performance programs to create clear expectation of measure performance.**

Section 5.3 Clustering Alternative

1. *Should CMS consider potential alternatives to k-means clustering in more detail? If so, what sort of change should CMS consider?*
2. *What other considerations should guide future CMS work regarding clustering?*

Vizient appreciates that CMS has made improvements in the k-means clustering approach. Vizient has generally supported these improvements; however, we would encourage more transparency in providing cluster analysis assessment statistics and validations, such as R-square, Pseudo F, CCC statistic, ANOVA, etc. – for researchers and statisticians to make fully informed recommendations on improving the methodology.

Section 5.4 Incorporation of Improvement

1. *Should CMS consider incorporating improvement in future iterations of the Overall Hospital Quality Star Rating?*
2. *What are conceptual benefits and risks of incorporating absolute score improvement into the Overall Hospital Quality Star Rating?*
3. *How should CMS operationalize this topic?*

Given the variability in scores generated using the latent variable modeling, the inclusion of an improvement score would not be helpful at this time. Additionally, Vizient believes that adding an improvement score would inherently introduce older data into the scoring, hurting the timeliness of the score. Thus, Vizient does not recommend including an improvement score, as doing so may lessen the impact of current performance.

Section 5.5 User-Customized Star Rating

1. *Should CMS consider introducing user-customization to the Overall Hospital Quality Star Rating?*
2. *What is the usability, utility, and validity of such a tool?*
3. *What are potential benefits and drawbacks to such a tool?*
4. *How could CMS incorporate such a tool into the existing Overall Hospital Quality Star Rating methodology?*

Vizient commends CMS for considering this innovative approach to place measure importance in the public's hands. However, given the Hospital Star Rating current complexity, Vizient sees challenges in the public's ability to understand the measures driving the ratings or finding measures that pertain to their particular needs or questions. Many of the measures used in the Star Ratings contain detailed, complicated algorithms that may be challenging for the average consumer to understand. Additionally, many of the main measures represented in the Star Ratings focus on limited clinical conditions such as heart failure, hip and knee replacements or COPD, which may not be the patient's specific condition or need.

A final step towards making the user-customized Star Ratings more informative to patients would be the inclusion of patient-reported outcome measures – which measure mobility, mental status and overall well-being. These measures compare providers based on questions that the average patient may find themselves wondering, such as 'how soon will I return to work?' or 'when can I go running again?' and answer the questions patients really want to know when they seek treatment. Unfortunately, our current health care measuring systems do not incorporate these measures and therefore miss an opportunity ripe for user-customization.

Conclusion

Vizient appreciates the opportunity to provide feedback on the Overall Hospital Quality Star Rating, and to inform the agency on how the methodology is impacting our members. We look forward to continuing to work with CMS to ensure patients and providers have access to reliable information. Vizient is encouraged that CMS has taken steps to seek additional input in order to deliver a better Star Ratings program, and looks forward to providing continued feedback and support.

In health care, patients expect reliable, consistent, high quality and scientifically based care to improve their quality of life. Health care providers expect the same when being measured for the care they deliver, while also seeking data and insights to drive continuous quality improvement. However, the current CMS Star Rating program falls short of these expectations by evaluating hospitals with methods, scoring incentives and data sets that do not portray an accurate or complete picture and include heterogeneous hospital comparisons which currently are misaligned with CMS' pay-for-performance programs.

Vizient supports CMS considering a more consistent weighting schema, for example as used in existing programs – while creating hospital cohorts that provide fair and meaningful performance evaluations. Additionally, Vizient strongly encourages CMS to explore leveraging more current data to provide more actionable and meaningful Star Ratings for performance improvement. We advocate for changes to the system that will support the core mission of the CMS Hospital Quality Star Rating of providing patients and the public with a clear, simple and objective mechanism for identifying top hospitals.

Vizient membership includes a wide variety of hospitals ranging from independent, community-based hospitals to large, integrated health care systems that serve acute and non-acute care needs. Additionally, many are specialized, including academic medical centers and pediatric facilities. Individually, our members are integral partners in their local communities, and many are ranked among the nation's top health care providers.

In closing, on behalf of Vizient, Inc., I would like to thank CMS for providing us this opportunity to comment. Please feel free to contact me at (202) 354-2600 or Chelsea Arnone, Director of Regulatory Affairs and Government Relations (chelsea.arnone@vizientinc.com), if you have any questions or if Vizient can provide any assistance as you consider these issues.

Respectfully submitted,

Shoshana Krilow
Vice President of Public Policy and Government Relations
Vizient, Inc.