American Hospital Association™
Advancing Health in America

Understanding CMS' Changes to Hospital Overall Star Ratings

Key Findings

- **CMS' new star ratings methodology preserves some year-to-year stability, but ratings remain volatile for hospitals reporting fewer measures.**
- **Under CMS' new peer grouping approach, 74% of hospitals are scored on all five measure groups with the remaining 26% split across the 14 combinations of measure groups.**
- **Star ratings remain volatile for hospitals reporting fewer measures, especially smaller, rural hospitals and critical access hospitals (CAHs).**
- **Ongoing systematic assessment of the methodology is vital, and CMS should evaluate potential approaches that make the basis of ratings more equitable.**

The Centers for Medicare & Medicaid Services (CMS) in 2020 overhauled the Hospital Overall Star Ratings methodology to make ratings more transparent, equitable, stable and predictable. To improve transparency the agency moved away from the statistically heavy and difficult to interpret latent variable modeling approach and instead employed a simple average of measure to calculate measure group scores. As a way to attempt to achieve more equitable comparisons, CMS also created peer groupings by number of reported measures. To improve the stability and predictability of star ratings the weights applied to the different measures would be set in advance by CMS. On the whole, while CMS expected the new methodology could change the overall ratings distribution, it believed period-to-period changes would remain fairly stable.

AHA commissioned KNG Health to assess whether these methodological changes have fully achieved CMS' goals, while also estimating expected performance variation and drivers of performance and variation under the new methodology.

The KNG analysis entailed:

- Comparing 2020 (old methodology) to 2021 (new methodology) hospital star ratings performance reported by CMS, assessing differences in ratings by hospital characteristics.
- Assessing the stability of CMS' new methodology by comparing hypothetical 2020 performance under the new methodology to 2021 actual performance; and
- Assessing the equitability of the ratings through analyses of peer groupings.

Overall the new methodology preserves some year-to-year stability. When applying the new methodology to both 2020 and 2021 the distribution of ratings are similar (see Figure 1). Sixty percent of hospitals could have expected to receive the same score in 2020, with 36% seeing an increase or decrease of one star rating.

American Hospital Association™
Advancing Health in America

Figure 1. Changes in Overall Star Ratings Under New Method, 2020 - 2021

STAR RATING	2020 COUNT	2020 % OF TOTAL	2021 COUNT	2021 % OF TOTAL	DIFFERENCE IN 2021
★	266	8.4%	187	5.9%	-79
★★	679	21.5%	657	20.8%	-22
★★★	1084	34.2%	955	30.2%	-129
★★★★	832	26.3%	934	29.5%	102
★★★★★	304	9.6%	432	13.6%	128

Totals only include those hospitals receiving a star rating

Source: KNG Health Consulting calculations using CMS reported performance data for 2020 and 2021.

In improving stability, however, star ratings remain volatile for hospitals reporting fewer measures, especially smaller, rural hospitals and critical access hospitals (CAHs). This volatility for hospitals reporting fewer measures is reflected in Figure 2, which applies CMS's old methodology to 2020 and the new methodology to 2021. Nearly one in four hospitals with less than 25% of reported measures could expect to see a change in star ratings by two or more stars.

Figure 2. Changes in Star Ratings By Percentage of Measures Reported, 2020 - 2021

PERCENTAGE OF MEASURES REPORTED	- TWO OR MORE STARS	- /+ ONE STAR	+ TWO OR MORE STARS
< 25%	15.4%	76.94%	7.7%
25 - 50%	8.8%	88.8%	2.4%
50 - 75%	7.5%	88.3%	4.1%
75 - 100%	1.8%	93.0%	5.1%

Source: KNG Health Consulting calculations using CMS reported performance data for 2020 and 2021.

Figure 3. Measure Reporting By Peer Group, 2021

# OF MEASURES REPORTED	COUNT	% OF TOTAL
FIVE	2464	73.8%
FOUR	544	16.3%
THREE	331	9.9%

Totals only include those hospitals receiving a star rating

Source: KNG Health Consulting calculations using CMS reported performance data for 2021.

Furthermore, the new peer grouping methodology introduced by CMS, which attempts to compare similar hospitals by number of measures reported, introduces an additional set of complexities. There are now 15 possible combinations of measure groups on which hospitals can be scored,

15 POSSIBLE COMBINATIONS OF MEASURE GROUPS

and while about 74% of hospitals are scored on all five measure groups, the remaining 26% are split across the rest of these combinations (see Figure 3). Furthermore, reporting within the same measure group does not mean that hospitals reported on the same measures. For example, two hospitals could each report three different readmission measures.

This means that a hospital's star rating may reflect a very different set of measure groups. This can be misleading and make it difficult to assess meaningful differences in quality between hospitals with different ratings. Even within measure groups there is significant variation in the number of measures on which hospitals are scored. For example, over 60% of CAHs reported less than half of the measures available for scoring in star ratings.

Additionally, CMS's new group and measure reweighting methodology also means the weight riding on a particular measure could vary from hospital to hospital. For example, certain measures, such as safety of care and patient experience, are more likely to not be reported (see Figure 4).

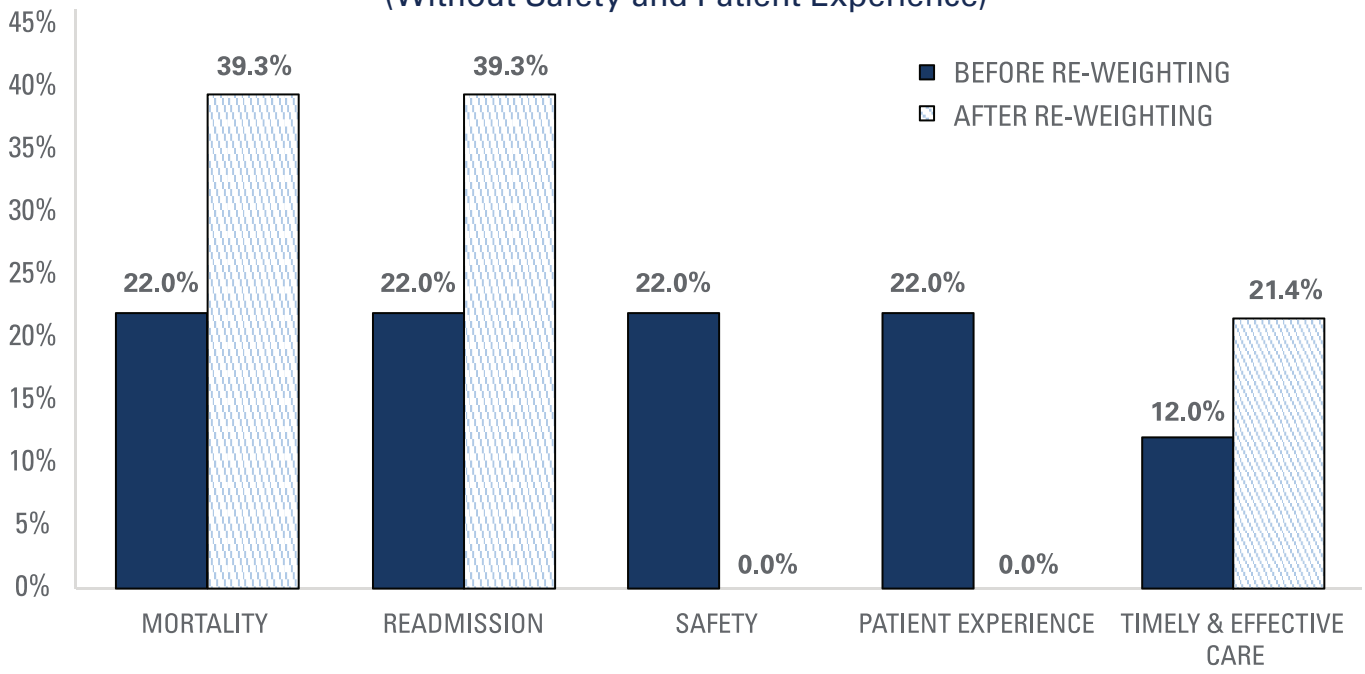
Figure 4. Percent of Hospitals Reporting By Measure Group, 2021

MEASURE GROUP	3 MEASURE PEER GROUP	4 MEASURE PEER GROUP	5 MEASURE PEER GROUP
MORTALITY	86.6%	89.5%	100.0%
SAFETY OF CARE	**13.1%**	**11.6%**	100.0%
READMISSION	89.2%	98.4%	100.0%
PATIENT EXPERIENCE	**20.2%**	96.6%	100.0%
TIMELY AND EFFECTIVE CARE	85.2%	97.5%	100.0%

Source: KNG Health Consulting calculations using CMS reported performance data for 2021.

For instance, when a hospital does not report a particular measure group, CMS rebalances the remaining measures reported by redistributing the weight of the excluded measure that is not reported. For example, if a hospital could not report on both the safety of care and patient experience measure groups, which each account for 22% of the rating score, the full 44% is redistributed across the other reported measures groups (see Figure 5). This rebalancing inflates the importance of the remaining measures.

Figure 5. New CMS Measure Re-weighting Example
(Without Safety and Patient Experience)



Note: This re-weighting example occurs when a hospital reports no measures within the safety and patient experience measure groups. Although these two measures are most commonly not reported, other redistributions are possible.

Implications

Given the new issues raised by CMS' latest changes to the Hospital Overall Star Ratings methodology, the AHA continues to be skeptical about whether quality rankings or rating scores can meaningfully reflect hospital performance without bias against particular types of hospitals. In these most recent changes to star ratings, we see that in attempting to solve one set of challenges around transparency and equitability, CMS has introduced additional challenges. The ratings are still not "apples-to-apples," as evidenced by the fact that there are so many combinations of groups and measures on which hospitals are assessed. As a result, the ratings are likely to remain unstable for hospitals reporting fewer measures.

Ongoing systematic assessment of the methodology is vital, and CMS should evaluate potential approaches that make the basis of ratings more equitable. For example, CMS should consider:

- Moving away from an overall rating, and instead explore approaches for scoring hospitals on individual topics.
- Assessing the peer grouping approach to see whether factors other than number of reported measure groups could be used, such as number of measures or CAH/Inpatient Prospective Payment System status.
- Ensuring public messaging around star ratings reflects that the rating is a function of the measures used in scoring as well as other methodological choices.

- Replacing unstable measures that examine rare events with more robust measures of safety and quality so that more hospitals have sufficient data to be compared on more measures.

Patients, consumers and other stakeholders should continue to assess ratings with considerable care, and engage their trusted health care providers in identifying specific measures that matter most to their care needs.

American Hospital Association™
Advancing Health in America

KNG HEALTH CONSULTING LLC

KNG supplied research and analytic support for this research brief.

www.aha.org
@AHAhospitals
@AHAadvocacy
facebook.com/AHAhospitals
@AHAhospitals
linkedin.com/company/american-hospital-association