Identifying performance gaps

To identify performance gaps in the GMCS, we computed the deciles of the measure scores using the EHR data.

Table VI.1 presents summary statistics (range, mean, minimum, maximum and decile distribution) of the GMCS measure scores across the 28 facilities, with Figure VI.1 illustrating the distribution of the scores. The average score was a 90.3%, with the lowest facility scoring at 83.7% and the highest facility at 98.2%. Despite the distribution of the GMCS measure scores being skewed towards high performance in our sample, there remains a substantial gap between the facilities in the best- (98%) and the worst-performing deciles (84%), indicating that many facilities have room for improvement.

**Table VI.1.** Distribution of the GMCS measure scores among 28 hospitals

|  | Mean Score | Number of facilities | Number of encounters |
| --- | --- | --- | --- |
| Overall | 90.3 | 28 | 145,846 |
| Minimum | 83.7 | 1 | 310 |
| Decile 1 | 84.0 | 3 | 4,945 |
| Decile 2 | 86.1 | 3 | 8,644 |
| Decile 3 | 87.9 | 3 | 10,299 |
| Decile 4 | 88.8 | 2 | 5,724 |
| Decile 5 | 89.5 | 3 | 11,824 |
| Decile 6 | 90.5 | 3 | 33,921 |
| Decile 7 | 91.8 | 2 | 19,748 |
| Decile 8 | 92.6 | 3 | 13,118 |
| Decile 9 | 94.2 | 3 | 9,877 |
| Decile 10 | 98.0 | 3 | 27,746 |
| Maximum | 98.2 | 1 | 6,304 |

Source: Mathematica analysis of the test-site data provided by Avalere Health and the Academy of Nutrition and Dietetics (N=28 facilities)

**Figure VI.1.** Histogram of facilities’ GMCS measure scores

A graph of distribution of gmcs

Description automatically generated

Source: Mathematica analysis of the test-site data provided by Avalere Health and the Academy of Nutrition and Dietetics (N=28 facilities)

We anticipate that the measure scores in our sample are skewed towards high performance as the facilities whose data was used for testing had been participating in the Malnutrition Quality-Improvement Initiative (MQii) initiative, which resulted in the improved malnutrition care workflow and the hospital's malnutrition diagnosis rate.

We expect that the distribution of measure scores in the universe of eligible facilities that are yet to integrate quality improvement into nutrition care will be skewed towards low to moderate performance. For example, in 2016, the University of Iowa hospital and clinics initiated a pilot program aimed at enhancing the workflow for malnutrition care and streamlining the process of evaluating inpatients, emphasizing improved communication and precise malnutrition assessment within their EHR. Prior to implementing the pilot, the institution estimated that malnutrition was diagnosed in less than 10% of inpatients. During the 4-month pilot period, malnutrition was identified in 42% of patients within these units, indicating a significant improvement compared to the previous diagnosis rate (Bechtold et al., 2023).

1. Topped out analysis

Given the generally high scores on the GMCS observed in our sample, we assessed whether the GMCS would be considered “topped out”. According to the Centers for Medicare & Medicaid Services (CMS), a measure is considered topped out when it has reached a plateau in performance, where further improvements are unlikely or minimal. In other words, it means that most providers have achieved high levels of performance, leaving little room for further improvement or differentiation between providers based on those measures.

While the topped-out analysis is not required for the MUC submission, this information can help the measure steward respond to possible questions about the distribution of the measure scores in the sample used for testing.

1. Methods

Intermediate outcome measures, such as the GMCS, are considered topped out if two criteria are met (CMS, 2014).

First, the truncated coefficient of variation (TCV)—a measure of relative variability of the scores in the sample—must be less than or equal to 0.1 or 10%. The TCV is calculated as the standard deviation divided by the mean of a truncated distribution of the measure scores, where facilities scoring in the top and bottom 5th percentiles of the measure have been removed from the data. A small TCV suggests that the distribution of individual measure scores is clustered tightly around the mean value, suggesting that it is not useful to draw distinctions between measure scores.

Second, the 75th and 90th percentiles of the measure must be statistically indistinguishable from one another. More specifically, the difference between the 75th and 90th percentiles must be less than or equal to twice the standard error (SE) of the 90th percentile. The SE of the 90th percentile can be computed using two approaches, as described below.

Criterion 1:

1. Remove the facilities with GMCS measure scores in the lower and upper 5th percentiles of the overall distribution of the measure scores. The purpose of this step is to remove the outliers that might disproportionately affecting the distribution of the measure scores.
2. Compute the mean and standard deviation in the truncated sample.
3. Compute the truncated coefficient of variation by dividing the standard deviation by the mean obtained in step c).

Criterion 2:

1. Create a bootstrap sample of the encounter-level data by randomly sampling encounters with replacement within each facility, keeping the original sample size within each facility.
2. Using the bootstrap sample from step a), calculate the GMCS measure score for each facility and record the 90th percentile of the measure scores within this sample.
3. Repeat steps a) and b) 1000 times, resulting in 1000 values of the 90th percentile of the GMCS measure scores.
4. Calculate the standard deviation of the 1000 values from step c).
5. Results

Table VI.3 presents the results of the topped-out analysis. It is important to stress that despite the generally narrow performance gap in the overall sample, the GMCS measure may have wider performance gaps within subsets of populations. Furthermore, it is important to keep in mind that GMCS measure consists of four distinct components. While Measure Observation (MO) 1 may indeed be topped out, affecting the overall measure status, the remaining components—MO2, MO3, and MO4—display notable variations, suggesting substantial potential for enhancement.

**Table VI.3.** Results of topped-out analysis of the GMCS

| Measure | Criterion 1 | | | Criterion 2 | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | TCV | Topped out based on Criterion 1? | N | 75th pctl. | 90th pctl. | Difference between 90th and 75th pctl. | 2 X SE for 90th pctl. | Topped out based on Criterion 2? |
| GMCS | 24\* | 0.04 | Yes (TCV ≤ 0.1) | 28 | 92.6 | 96.0 | 3.4 | 0.4 | No (3.4 > 0.4) |

Source: Mathematica analysis of the test-site data provided by Avalere Health and the Academy of Nutrition and Dietetics (N=28 facilities)

\*Note: Under the TCV approach, facilities with GMCS scores <=5th percentile and >=95th percentile are excluded from calculations (N=4)

1. Performance categorization

To further demonstrate performance gaps in the GMCS, we categorized each facility according to whether their performance on the GMCS is better, worse, or no different from average performance.

1. Methods

First, we computed a 95% confidence interval (C.I.) around each facility’s GMCS measure score as:

95% C.I. for a facility GMCS score *=*

where *p* is a GMCS score, *σ* is the standard deviation of Measure Observation 6 scores for a given facility, and *n* is the number of encounters in that facility.

Next, we determined whether the 95% C.I. for each facility’s GMCS score included the mean GMCS score in our sample (90.3%; the score corresponding to the performance of an average facility) and assigned each facility to one of three performance categories: facilities that perform better (C.I. lies above the mean), worse (C.I. lies below the mean), or no different from an average facility (C.I. contains the mean) in our sample.

1. Results

Table VI.4 shows the performance categories for facilities based on the GMCS measure. A total of 12 facilities (42.9%) performed “Better than average,” 6 facilities (21.4%) performed “No different than average,” and the remaining 10 facilities (35.7%) performed “Worse than average.” These findings suggest that despite a relatively narrow performance gap, there is room for quality improvement even among our sample of high-performing hospitals, as 16 (57.1%) of facilities performed either worse than or not significantly different from the average facility in the data.

**Table VI.4.** Breakdown of facility performance categories

| Performance group | Number of facilities | Percent of facilities |
| --- | --- | --- |
| Better than average | 12 | 42.9 |
| Not significantly different from the average | 6 | 21.4 |
| Worse than average | 10 | 35.7 |
| Total | 28 | 100 |

Source: Mathematica analysis of the test-site data provided by Avalere Health and the Academy of Nutrition and Dietetics (N=28 facilities)

References

Bechtold, Matthew L., et al. "Impact of a nutrition support therapy on hospital‐acquired infections: a value analysis." *Nutrition in Clinical Practice* 36.5 (2021): 1034-1040.