

**Overall Hospital Quality Star Rating on *Hospital Compare*  
December 2017 Updates and Specifications Report**

December 2017

## **Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE) Project Team**

**Arjun K. Venkatesh, MD, MBA, MHS\*** – Project Lead

**Susannah M. Bernheim, MD, MHS** – Project Director

**Li Qin, PhD** – Lead Analyst

**Haikun Bao, PhD** – Supporting Analyst

**Jaymie Simoes, MPH** – Project Manager

**Erica Norton, BS** – Research Associate

**Megan Wing, MA** – Research Associate

**Grace Glennon, MS** – Research Associate

**Rushi Shah, BS** – Research Assistant II

**Jeph Herrin, PhD\*** – Statistical Consultant

**Haiqun Lin, MD, PhD** – Statistical Consultant

**Zhenqiu Lin, PhD** – Analytics Director

**Harlan M. Krumholz, MD, SM\*** – Principal Investigator

\*Yale School of Medicine

### **Acknowledgements**

This work is a collaborative effort, and the authors gratefully acknowledge and thank our many colleagues and collaborators for their thoughtful and instructive input. Acknowledgment of input does not imply endorsement of the methodology and policy decisions.

We would like to acknowledge and thank Lein Han, PhD, Vinitha Meyyur, PhD, Golden Davis, MS, Reena Duseja, MD, Kristie Baus, MS, RN, William Lehrman, MA, PhD, Wilfred Agbenyikey, ScD, MPH, and Cynthia Tourison at the Centers for Medicare & Medicaid Services (CMS) for their contributions to this work.

We would also like to acknowledge and thank the contribution and work from Kit Cooper, BS and Dawn Beard, LPN from the Lantana Consulting Group, as well as Rachel Singer, PhD, MPH, MPA from NORC at the University of Chicago.

Additionally, we would like to acknowledge Hospital Consumer Assessment of Healthcare Providers Systems (HCAHPS), Centers for Disease Control & Prevention (CDC), Mathematica Policy Research (MPR), and a statistical consultant from Harvard University, Sharon-Lise Normand, PhD.

We would also like to thank Lisa Suter, MD, Elizabeth Drye, MD, SM, Lori Geary, MPH, Angela Hsieh, PhD, and Tamara Mohammed, MHA, CHE from CORE.

Furthermore, we would like to thank the National Partnership for Women and Families (NPWF) for their collaboration in improving patient and consumer outreach and involvement.

Finally, we would like to recognize the commitment and contributions of the members of our Technical Expert Panel (TEP), our Patient & Patient Advocate Work Group, and our Provider Leadership Work Group (PLWG).

# 1. How to Use This Report

Under contract with the Centers for Medicare & Medicaid Services (CMS), Yale New Haven Health Services Corporation – Center for Outcomes Research & Evaluation (CORE) has developed the methodology for the Overall Hospital Quality Star Rating, summarizing the quality information conveyed by many of the measures publicly reported on *Hospital Compare*. The purpose of this report is to provide an overview of the methodology for calculating the Star Rating (v2.0) and provide updated national results for the December 2017 *Hospital Compare* release. A more detailed version of the methodology and process for developing the Star Rating (“Overall Hospital Quality Star Rating v2.0”) can be found on *QualityNet* at [www.qualitynet.org](http://www.qualitynet.org) > [Hospitals-Inpatient](#) > [Overall Hospital Ratings](#). Questions regarding the Overall Hospital Quality Star Rating can be sent to [cmsstarratings@lantanagroup.com](mailto:cmsstarratings@lantanagroup.com).

This Overall Hospital Quality Star Rating Updates and Specifications Report (December 2017) is organized into the following sections:

- Section 2: Objective of Overall Hospital Quality Star Rating
- Section 3: Overall Hospital Quality Star Rating Methodology
  - 3.1. [Summary of Updates to the Methodology](#)
  - 3.2. [Overview of Five Steps of Star Rating Methodology](#)
  - 3.3. [Step 1: Selection of Measures for Inclusion in Star Rating](#)
  - 3.4. [Step 2: Assignment of Measures to Groups](#)
  - 3.5. [Step 3: Calculation of Latent Variable Model Group Scores](#)
  - 3.6. [Step 4: Calculation of the Overall Hospital Summary Score as a Weighted Average of Group Score](#)
  - 3.7. [Step 5: Application of Minimum Thresholds for Receiving a Star Rating](#)
  - 3.8. [Step 6: Application of Clustering Algorithm to Obtain a Star Rating](#)
- Section 4: Results for December 2017 Implementation of Star Ratings
  - 4.1. [Group Performance Category](#)
  - 4.2. [Distribution of Star Rating and Group Performance Categories](#)
- Appendix A: [Flowchart of the Five-Step Overall Hospital Quality Star Rating Methodology](#)
- Appendix B: [Measures Excluded from the December 2017 Star Rating \(N=65\) by Exclusion Criterion](#)
- Appendix C: [Measures Included in the December 2017 Star Rating \(N=57\) by Group](#)
- Appendix D: [Measure Loadings by Group for December 2017](#)
- Appendix E: [Comparative Analysis of the Updated Star Ratings Methodology](#)

## **2. Objective of Overall Hospital Quality Star Rating**

The primary objective of the Overall Hospital Quality Star Rating project is to summarize information from the existing measures on *Hospital Compare* in a way that is useful and easy to interpret for patients and consumers through the development of a statistically sound methodology. Consistent with other CMS Star Rating programs, this methodology assigns each hospital between one and five stars, reflecting the hospital's overall performance on selected quality measures.

The Overall Hospital Quality Star Rating is designed to provide summary information for consumers about existing publicly-reported quality data. In the case of *Hospital Compare*, the Overall Hospital Quality Star Rating will be complimentary to existing efforts, such as the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) star ratings (implemented in April 2015), and will not replace the reporting of any individual quality measures. In what follows, "Star Rating(s)" refers to the Overall Hospital Quality Star Rating unless otherwise noted.

### 3. Overall Hospital Quality Star Rating Methodology

#### 3.1. Summary of Updates to the December 2017 Star Rating Calculations

The Overall Star Rating methodology for December 2017 includes four updates that have been developed through collaboration with stakeholders. The enhancements to the methodology include:

Update	Description	Rationale
1. Combining non-adaptive quadrature with adaptive quadrature	Adaptive quadrature utilizes prior calculations at each calculation iteration of the statistical models to improve stability of measure group score estimates.	The addition of adaptive quadrature calculates a solution, or convergence, of the statistical model with higher accuracy than non-adaptive quadrature alone.
2. Utilizing multiple iterations of k-means clustering to complete convergence	This replaces the previous SAS procedure which involved one iteration of K-means clustering used to assign hospitals to 5 star categories.	Running k-means clustering to complete convergence provides more reliable and stable star ratings assignments.
3. Resequencing of reporting thresholds	This method applies the reporting thresholds before the k-means clustering instead of after.	Since k-means clustering is a comparative analysis, limiting this calculation to hospitals that meet reporting requirements is reasonable.
4. Removal of hospital summary score winsorization	The previous methodology winsorized, or trimmed extreme outlier hospital summary scores, to promote a broader distribution of star ratings by minimizing the effect of outlier hospitals on clustering	Given that use of complete convergence for clustering results in a broader distribution of star ratings, this winsorization step is no longer necessary.

Taken together, these enhancements have been implemented to improve the Overall Star Rating methodology. Prior to making such changes to the methodology, CMS sought stakeholder feedback through a Technical Expert Panel (TEP), two Work Groups, and a public comment period. CMS reflected on these recommendations to anticipate any unexpected results.

## 3.2. Overview of Six Steps of the Star Rating Methodology

The methodology to calculate the Star Rating is comprised of a six-step process. These steps are listed below and are described in greater detail in subsequent sections (see [Appendix A](#)).

- Step 1: Selection and standardization of measures for inclusion in the Star Rating
- Step 2: Assignment of measures to groups
- Step 3: Calculation of latent variable model group scores
- Step 4: Calculation of hospital summary scores as a weighted average of group scores
- Step 5: Application of public reporting thresholds for receiving a Star Rating
- Step 6: Application of clustering algorithm to translate a summary score into a Star Rating

The measures were first selected based on their relevance and importance as determined through stakeholder and expert feedback. The selected measures were standardized to be consistent in terms of direction and magnitude, with outlying values trimmed (Step 1). In Step 2, the measures were organized into seven groups by measure type. In Step 3, the standardized measures for each group were used to construct a latent variable statistical model that reflected the dimension of quality represented by the measures within the given group. Each of the seven statistical models generated a hospital-specific group score, which is obtained as a prediction of the latent variable. The term prediction is used to represent the realized value of the latent variable. In Step 4, a weight was applied to each group score, and all available groups were averaged to calculate a hospital summary score. In Step 5, the minimum public reporting thresholds for receiving a Star Rating are applied to determine inclusion or exclusion of hospitals from Star Rating calculations. Finally, in Step 6, hospital summary scores were organized into five categories using a clustering algorithm to assign a Star Rating.

Of note, CMS also reports hospital performance at the group level, separately categorizing each of a hospital's available group scores into one of three group performance categories (above, same as, and below the national average). These performance categories provide additional detail to patients and consumers using *Hospital Compare* (see [Section 4.2](#)).

### **3.3. Step 1: Selection and Standardization of Measures for Inclusion in the Star Rating**

#### ***Criteria for Selecting Measures***

CMS determined and vetted measure selection criteria with stakeholders through the Technical Expert Panel (TEP) and opportunities for public input to ensure that the Star Rating captured the diverse aspects of quality represented by the measures on *Hospital Compare*.

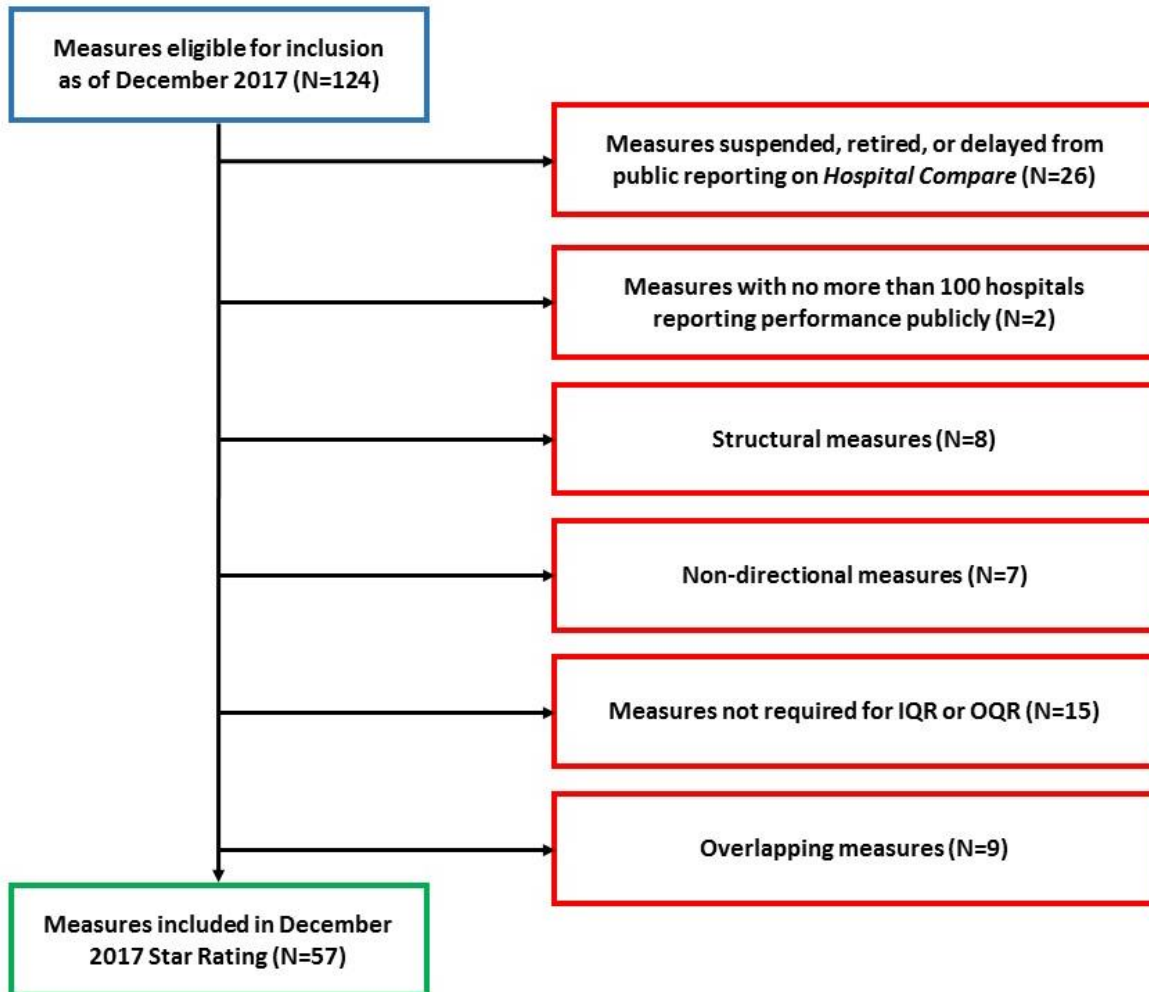
Because the Star Rating is intended for acute care hospitals, CMS first excluded all measures on *Hospital Compare* that were specific to specialty hospitals (such as cancer hospitals or inpatient psychiatric facilities), or ambulatory surgical centers prior to applying any measure selection criteria. With these measures omitted, the total number of measures eligible for inclusion in the Star Rating for December 2017 is 124 measures. The Star Rating methodology further limited the number of measures for inclusion in order to maintain a sound methodology through measure selection criteria, which are presented in the subsequent text and in [Figure 1](#).

#### ***Measure Selection Criteria***

CMS uses the following criteria to exclude measures from the Star Rating calculation:

1. Measures suspended, retired, or delayed from public reporting on *Hospital Compare*;
2. Measures with no more than 100 hospitals reporting performance publicly;
3. Structural measures;
4. Measures for which it is unclear whether a higher or lower score is better (non-directional);
5. Measures not required for Inpatient Quality Reporting (IQR) Program or Outpatient Quality Reporting (OQR) Program; and
6. Overlapping measures (for example, measures that are identical to another measure, or measures with substantial overlap in cohort and/or outcome).

Figure 1. Measure Selection Flowchart (December 2017)



For a complete list of the measures excluded in December 2017, please refer to [Appendix B](#).

### ***Standardization of Measure Scores***

For all selected measures, CMS transforms the measures into a single, common scale to account for differences in measure score format, differences in score direction, and the occurrence of extreme outliers. A measure is standardized by subtracting the national mean of measure scores from a hospital's measure score and dividing it by the standard deviation across hospitals. Measure direction is standardized by reversing the direction of the standardized scores for all measures for which "lower score is better," and converting them into "higher score is better" measures. Finally, CMS utilizes Winsorization to limit the influence of measures with extreme outlier values at the 0.125<sup>th</sup> percentile ( $Z=-3$ ) and the 99.875<sup>th</sup> percentile ( $Z=3$ ). Winsorization is a common strategy used to set extreme outliers to a specified percentile of the data. All standardized measure scores above 3 are set to be 3, and all standardized below -3 are set to be -3.



### **3.4. Step 2: Assignment of Measures to Groups**

#### ***Approach to Grouping Measures***

CMS organizes measures into groups by measure type. Group names were finalized with input from a Patient & Patient Advocate Work Group, convened in collaboration with the National Partnership for Women and Families (NPWF), and previous CMS consumer testing. The Star Rating groups are:

- Mortality
- Safety of Care
- Readmission
- Patient Experience
- Effectiveness of Care
- Timeliness of Care
- Efficient Use of Medical Imaging

#### ***Measures by Group for December 2017***

We assigned each measure included in the Star Rating to one of seven mutually exclusive groups: Mortality (N=7), Safety of Care (N=8), Readmission (N=9), Patient Experience (N=11), Effectiveness of Care (N=10), Timeliness of Care (N=7), and Efficient Use of Medical Imaging (N=5). For a complete list of the measures in each group, please refer to [Appendix C](#).

### **3.5. Step 3: Calculation of Latent Variable Model Group Scores**

#### ***Overview of Latent Variable Model (LVM)***

CMS employs latent variable modeling (LVM) to estimate a group score for the dimension of quality represented by the measures in each group. LVM is a statistical modeling approach that has been used to summarize information in a variety of settings ranging from education to healthcare. For the Star Rating, LVM assumes each measure reflects information about an underlying, unobserved dimension of quality. A separate LVM is constructed for each group so that a total of seven LVMs are used to calculate the Star Rating. The LVM accounts for the relationship, or correlation, between measures for a single hospital. Measures that are more consistent with each other, as well as measures with larger denominators, have a greater influence on the derived latent variable. Each model estimates, for each hospital, the value of a single latent variable representing the underlying, unobserved dimension of quality; this estimate is the hospital's group score.

#### ***Update for December 2017***

The current estimation process involves numerical calculation of an integral for which a quadrature technique is used. Previous Overall Star Ratings methodology uses a specific version of quadrature called non-adaptive Gaussian Quadrature. This approach is being updated for

December 2017 public reporting to utilize adaptive quadrature. Adaptive quadrature utilizes prior calculations at each iteration to strategically place quadrature points over the LVM integral range so fewer quadrature points are needed to achieve higher accuracy. The updated methodology first utilizes non-adaptive Gaussian Quadrature to generate initial values that are subsequently used to estimate each group LVM using adaptive quadrature. This approach is found to generate stable hospital estimates with fewer quadrature points. Also, many studies of comparisons suggested that adaptive quadrature should be used for its balance of computational efficiency and accuracy.<sup>1-3</sup>

### ***Loadings of Measures within Each Group***

As noted above, measures that are more consistent, or more correlated, with other measures within the group have a greater influence on the hospital's group score. The influence of an individual measure on the group score is represented by the measure's "loading."

A loading is empirically derived for each measure in a group when applying the LVM; these statistically estimated measure loadings are regression coefficients based on maximum likelihood methods using observed data and are not subjectively assigned. A loading reflects the degree of the measure's influence on the group score relative to the other measures included in the same group. A measure's loading is the same across all hospitals. Measures with higher loadings are more strongly associated with the group score and the other measures within that group. All measures included in the Star Rating have an effect on the group score; however, measures with higher loadings have a greater association (or impact) on the group score than measures with substantially lower loadings. The loadings for the December 2017 Star Rating are reported in [Appendix D](#).

Please note, the loadings for an individual measure are re-estimated each time the Star Rating is updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

### ***Accounting for Measure Sampling Variation***

Hospitals' reported measure scores may include different numbers of patients, depending on the measure. For each measure, some hospitals may report a score based on data from fewer cases while other hospitals report scores based on more cases, resulting in differing precision for each hospital's individual measure score. This variability in precision is usually known as "sampling variation."

CMS gives more weight to measure scores that are more precise by using a weighted likelihood method. This method uses the hospital's measure denominator (hospital case count or sample size) to weight the observed value. A weighted likelihood ensures that a hospital with a larger denominator, or a more precise measure score, contributes more in calculating the loadings used to estimate the group score.

### 3.6. Step 4: Calculation of the Hospital Summary Score as a Weighted Average of Group Scores

#### *Weighting Scheme*

The seven group scores are combined to create a hospital summary score. To calculate the hospital summary scores, CMS takes a weighted average of that hospital’s available group scores, assigning each group a policy-based weight ([Table 1](#)). This weighting scheme was modified from that used for the Hospital Value-based Purchasing (VBP) program. Additionally, these weights were thoroughly vetted through the TEP, opportunities for public input, hospital dry run, and the Patient & Patient Advocate Work Group.

**Table 1. Measure Group Weight for the Star Rating**

Measure Group	Star Rating Weights
Mortality (N=7)	22%
Safety of Care (N=8)	22%
Readmission (N=9)	22%
Patient Experience (N=11)	22%
Effectiveness of Care (N=10)	4%
Timeliness of Care (N=7)	4%
Efficient Use of Medical Imaging (N=5)	4%

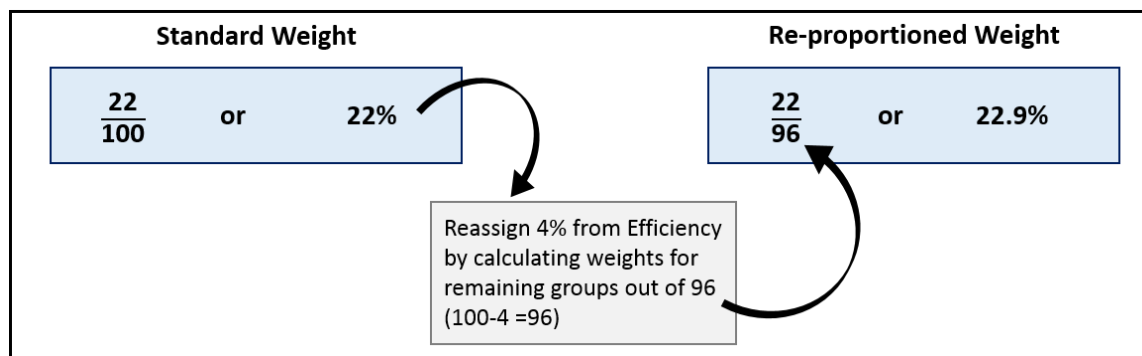
#### *Method for Re-Weighting When Missing Group(s)*

If a hospital reports no measures for a given measure group, that group is considered to be “missing.” When a hospital is missing one or more groups, CMS applies the same approach as the Hospital VBP program, re-proportioning the weight of the missing group(s) across the groups for which the hospital does report measures. [Table 2](#) and [Figure 2](#) provide examples of how CMS adjusts the weighting scheme for a hospital that is missing the Efficient Use of Medical Imaging group.

**Table 2. Example Re-Weighting Scheme when Missing Efficient Use of Medical Imaging Group**

Measure Group	Standard Weight	Re-proportioned Weight
Mortality	22%	22.9%
Safety of Care	22%	22.9%
Readmission	22%	22.9%
Patient Experience	22%	22.9%
Effectiveness of Care	4%	4.2%
Timeliness of Care	4%	4.2%
Efficient Use of Medical Imaging (N=0)	4%	---

Figure 2. Example Calculation for Re-Proportioning Group Weights



### ***Winsorization of Summary Scores***

Prior to December 2017, CMS performed a Winsorization of hospital summary scores. Winsorization is a common strategy used to set extreme outliers to a specified percentile of the data and was used to limit the influence of outliers and broaden the distribution of star ratings.

#### Update for December 2017

This Winsorization step is being removed in December 2017. The use of complete convergence for k-means clustering results in a broader distribution of star ratings making winsorization no longer necessary.

### **3.7. Step 5: Application of Minimum Thresholds for Receiving a Star Rating**

#### ***Minimum Thresholds for Receiving a Star Rating***

CMS evaluated and developed standards regarding the minimum number of measures and groups a hospital must report in order to receive a publicly reported Star Rating on *Hospital Compare*. CMS set these thresholds to allow for as many hospitals as possible to receive a Star Rating without sacrificing the validity and reliability of the Star Rating methodology.

On average, hospitals reported 5 groups and 36 measures for December 2017. A total of 3,692 (80.6%) hospitals on *Hospital Compare* met the public reporting threshold for receiving a Star Rating for December 2017.

Please note results included in this report may differ from the results posted on *Hospital Compare* due to data suppressed by CMS for one or more quarters. CMS may suppress data for various reasons, like data inaccuracies.

#### ***Minimum Threshold of Measures per Group***

The minimum measure threshold for December 2017 is three measures per group.

### ***Minimum Threshold of Groups in Summary Score***

The minimum group threshold for December 2017 is three groups with at least one outcome group (that is, Mortality, Safety of Care, or Readmission).

If a hospital met the minimum threshold of having three groups (one of which must be an outcome group) with at least three measures in each of the three groups, any other measures reported by the hospital were also included in the hospital's Star Rating. That is, any additional measures were included even if the hospital did not meet the minimum three measure threshold for a given group. This decision ensured that the Star Rating was inclusive of publicly reported measures and was vetted with the public through the second opportunity for public input.

### ***Reporting Thresholds for December 2017 Implementation***

#### ***Update for December 2017***

In previous reporting periods, the reporting threshold was applied after all other calculations resulting in inclusion of all hospitals in the estimation of group LVMs, the calculation of hospital summary scores and the application of k-means clustering. This sequence was originally designed to maximize the amount of "information" available for clustering hospitals and to apply the public reporting threshold in a manner that could potentially provide hospitals with ratings information privately if the number or type of measures reported were not sufficient to receive a star rating publicly. This approach mirrors the approach used for many individual quality measures reported on Hospital Compare.

Based on prior stakeholder feedback and regularly planned analyses, the application of the reporting threshold has been re-sequenced to occur prior to clustering. In other words, only those hospitals that meet the public reporting thresholds would be clustered into star ratings. Because k-means clustering is inherently a comparative analytic procedure, there may be better conceptual basis to apply clustering only to the subset of hospitals for which a star rating is ultimately to be reported.

## **3.8. Step 6: Application of Clustering Algorithm to Obtain a Star Rating**

### ***Approach for Translating a Summary Score to a Star Rating***

To translate each hospital's summary score to a rating between one and five stars, CMS applies k-means clustering.

### ***Overview of k-Means Clustering***

The k-means clustering analysis is a standard method for creating categories (or clusters) so that observations in each category are closer to their category mean than to any other category means. The number of categories is pre-specified; CMS specifies five categories, so that the k-means clustering analysis generates five categories (clusters) based on hospital summary scores in a way

that minimizes the distance between summary scores (observations) and the middle value of their assigned cluster (category mean). It organizes hospitals into one of five categories such that a hospital's summary score is "more like" that of the other hospitals in the same category and "less like" the summary scores of hospitals in the other categories. The Star Rating categories are structured such that the lowest group is one star and the highest group is five stars.

Update for December 2017

K-means clustering is used in the Overall Star Rating methodology in order to create five star categories and assign hospitals to star categories in a way that ensures hospital summary scores are more similar within that star category than in other star categories. Through stakeholder input, and as a part of the planned Star Rating reevaluation, we have incorporated the use of multiple iterations of clustering to achieve complete convergence as a technical enhancement to the methodology. Iterating the k-means clustering process until fewer and fewer hospitals shift star ratings categories provides more reliable and stable star ratings assignments. The previous SAS software procedure utilized the recommended default setting, which runs a single iteration to identify the k-means clusters or star categories.

## 4. Results for December 2017 Implementation of Star Ratings

### *Group Performance Category*

In addition to a hospital’s Star Rating, CMS reports a group performance category for each of a hospital’s available (in other words, meeting the minimum threshold) measure groups. To calculate a performance category, a hospital’s group score is compared to the national average group score. The LVM for each group produces a point estimate and standard error that can be used to construct a 95% confidence interval for each hospital’s group score for comparison to the national mean group score. The group performance categories are:

- “Above the national average,” defined as a group score with a confidence interval that falls entirely *above* the national average;
- “Same as the national average,” defined as a group score with a confidence interval that *includes* the national average; and
- “Below the national average,” defined as a group score with a confidence interval that falls entirely *below* the national average.

### *Distribution of the Star Rating and Group Performance Categories*

The Star Rating for December 2017 public reporting is calculated using December 2017 *Hospital Compare* data. The frequency of hospitals by each Star Rating category is shown in [Table 3](#). Of note, the minimum and maximum score for each category will change with each reporting period based on the underlying distribution of hospital summary scores.

**Table 3. Frequency of Hospitals by Star Category using k-Means**

Rating	Number of Hospitals (Frequency)	Summary Score Range in Cluster	Mean (sd)
1 Star	260 (7.0%)	-2.12,-0.78	-1.09 (0.27)
2 Star	753 (20.4%)	-0.77,-0.26	-0.47 (0.14)
3 Star	1,187 (32.2%)	-0.26,0.13	-0.05 (0.11)
4 Star	1,155 (31.3%)	0.13,0.56	0.31 (0.12)
5 Star	337 (9.1%)	0.56,2.21	0.83 (0.26)

Note: The total number of hospitals in the *Hospital Compare* dataset as of December 2017 is 4,579 hospitals.

Results shown are for all hospitals meeting the reporting criteria (N=3,692).

Please note results included in this report may differ from the results posted on *Hospital Compare* due to data suppressed by CMS for one or more quarters. CMS may suppress data for various reasons, like data inaccuracies.

[Table 4](#) displays the frequency of hospitals in each performance category by group.

**Table 4. Frequency of Hospitals by Group Performance Category**

<b>Group</b>	<b>Above the National Average</b>	<b>Same as the National Average</b>	<b>Below the National Average</b>
Mortality (N=7)	389 (11.3%)	2,703 (78.5%)	352 (10.2%)
Safety of Care (N=8)	1,148 (43.4%)	584 (22.1%)	912 (34.5%)
Readmission (N=9)	1,537 (39.5%)	977 (25.1%)	1,378 (35.4%)
Patient Experience (N=11)	1,213 (34.8%)	1,152 (33.1%)	1,120 (32.1%)
Effectiveness of Care (N=10)	70 (1.9%)	3,421 (91.5%)	250 (6.7%)
Timeliness of Care (N=7)	1,159 (31.4%)	1,533 (41.6%)	996 (27.0%)
Efficient Use of Medical Imaging (N=5)	502 (17.2%)	1,988 (68.0%)	435 (14.9%)

Note: The total number of hospitals in the *Hospital Compare* dataset as of December 2017 is 4,579 hospitals.

Results shown are for all hospitals meeting the reporting criteria (N=3,692).

Please note results included in this report may differ from the results posted on *Hospital Compare* due to data suppressed by CMS for one or more quarters. CMS may suppress data for various reasons, like data inaccuracies.



# Appendix A: Flowchart of the Six-Step Overall Hospital Quality Star Rating Methodology

Figure A.2 The Six Steps of the Current Star Rating Methodology

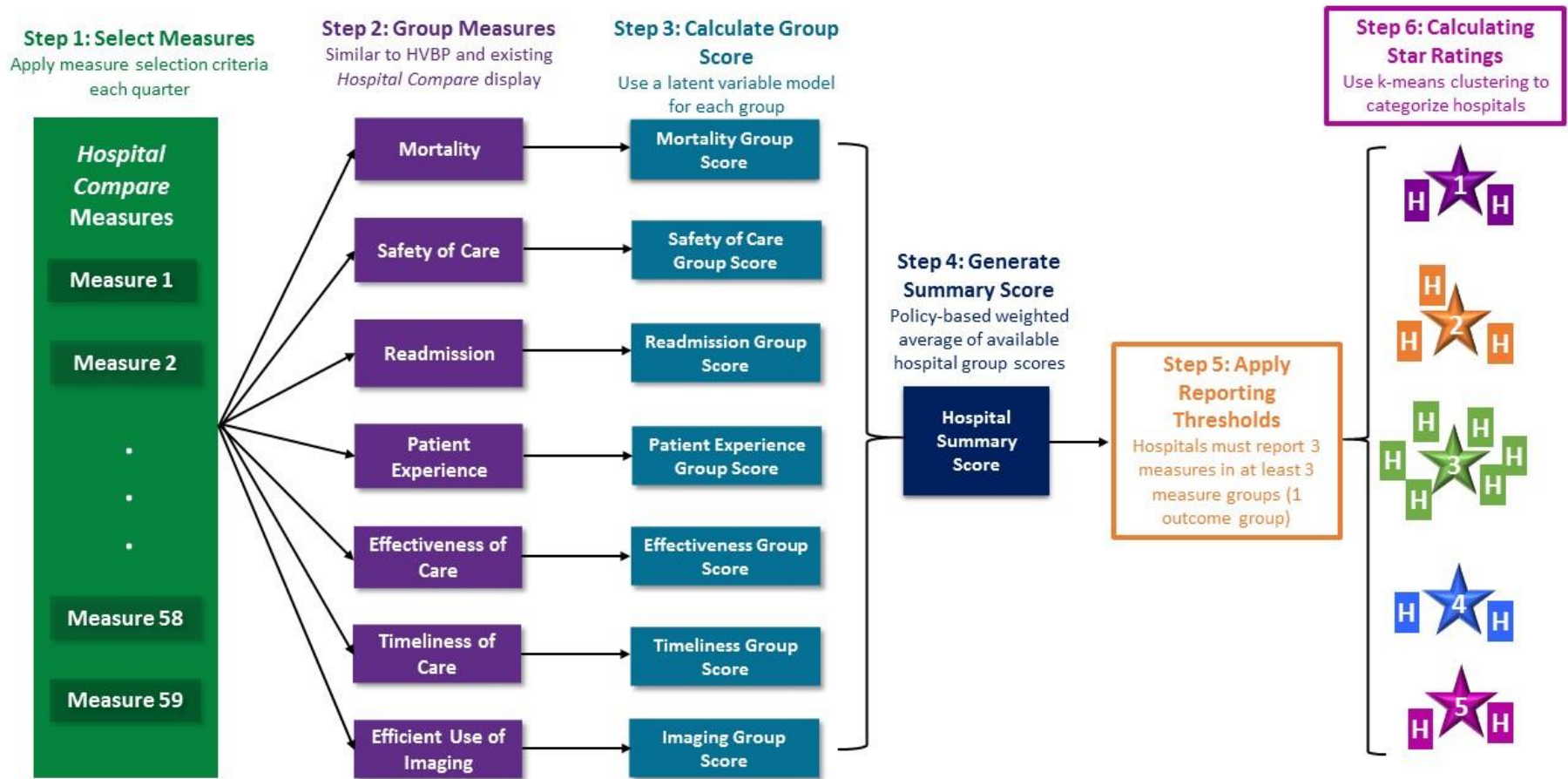
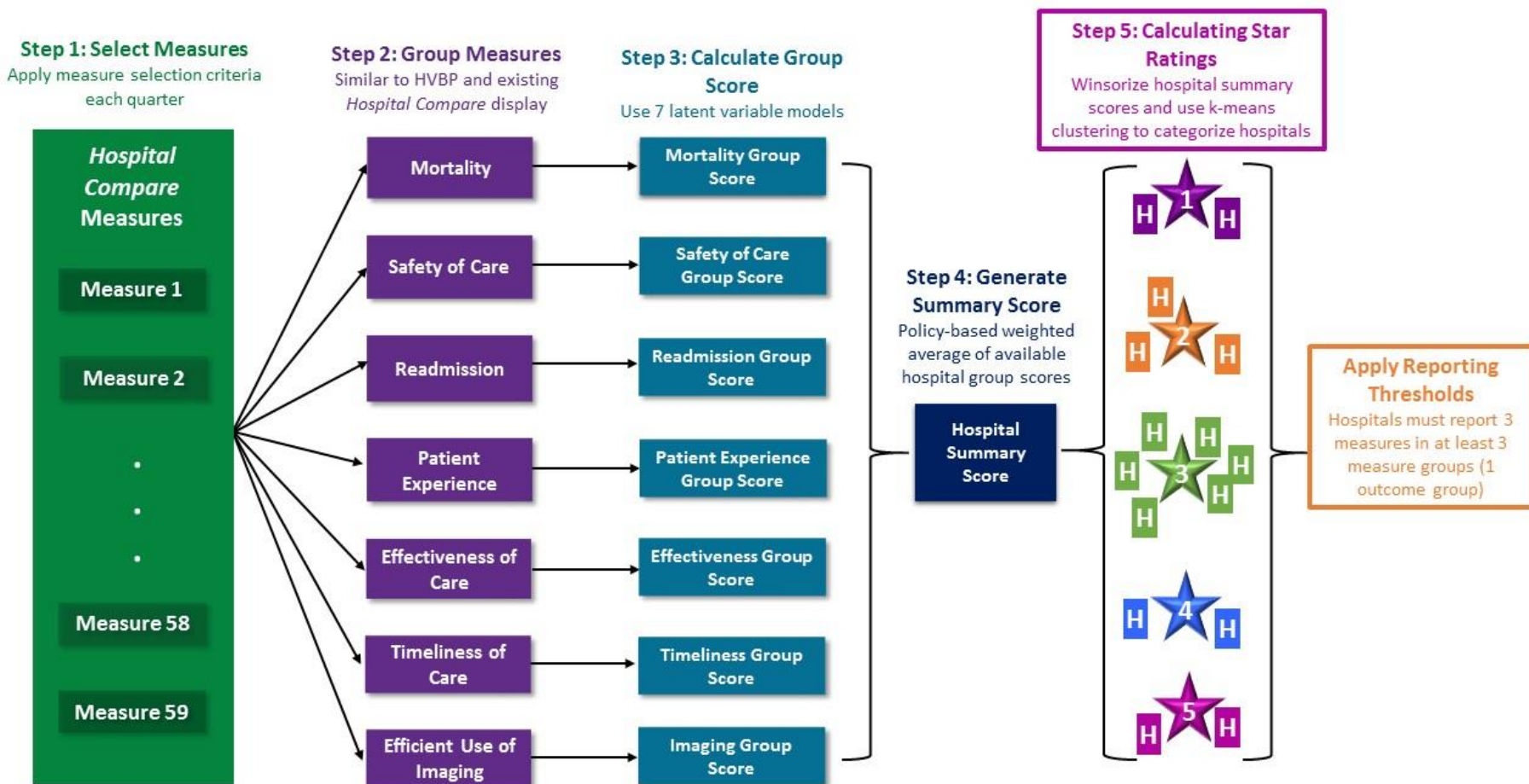


Figure A.1. The Five Steps of the Previous Star Rating Methodology



## **Appendix B: Measures Excluded from December 2017 Star Rating (N=67) by Exclusion Criterion**

### **Measures suspended, retired, or delayed from public reporting**

1. AMI-2 Aspirin Prescribed at Discharge
2. AMI-10 Statin Prescribed at Discharge
3. CAC-1 Relievers for Inpatient Asthma
4. CAC-2 Systemic Corticosteroids for Inpatient Asthma
5. HF- 1 Discharge Instructions
6. HF-3 ACEI or ARB for LVSD
7. OP-6 Timing of Antibiotic Prophylaxis
8. OP-7 Prophylactic Antibiotic Selection for Surgical Patients
9. PN-3b Blood Cultures Performed in the ED prior to Initial Antibiotic Received in Hospital
10. SCIP-Inf-4 Cardiac Surgery Patients with Controlled Postoperative Blood Glucose
11. SCIP-Inf-10 Surgery Patients with Perioperative Temperature Management
12. SM-PART-STROKE Participation in a Systematic Clinical Database Registry for Stroke Care
13. MV Number of Medicare Patient Discharges for Selected MS-DRGs
14. AMI-7a Fibrinolytic Therapy Received Within 30 Minutes of Hospital Arrival
15. SM-PART-CARD Participation in a Systematic Clinical Database Registry for Cardiac Surgery
16. VTE-1 Venous Thromboembolism Prophylaxis
17. VTE-2 Intensive Care Unit Venous Thromboembolism Prophylaxis
18. VTE-3 Venous Thromboembolism Patients with Anticoagulation Overlap Therapy
19. VTE-5 Venous Thromboembolism Warfarin Therapy Discharge Instructions
20. CAC-3 Home Management Plan of Care (HMPC) Document Given to Patient/Caregiver
21. STK-1 Venous Thromboembolism (VTE) Prophylaxis
22. STK-4 Thrombolytic Therapy
23. STK-6 Discharged on Statin Medication
24. STK-8 Stroke Education
25. HAI-1a Central Line-Associated Bloodstream Infection (CLABSI) – ICU Only
26. HAI-2a Catheter-Associated Urinary Tract Infection (CAUTI) – ICU Only

### **Measures with less than or equal to 100 hospitals reporting**

1. OP-1 Median Time to Fibrinolysis
2. OP-2 Fibrinolytic Therapy Received Within 30 Minutes of Emergency Department Arrival

### **Structural measures without evidence of an association with changes in clinical practice or improved outcomes**

1. ACS-REGISTRY Participation in a Multispecialty Surgical Registry

2. SM-PART-GEN-SURG Participation in a Systematic Clinical Database Registry for General Surgery
3. SM-PART-NURSE Participation in a Systematic Clinical Database Registry for Nursing Sensitive Care
4. OP-12 The Ability for Providers with HIT to Receive Laboratory Data Electronically Directly into their ONC-Certified EHR System as Discrete Searchable Data
5. OP-17 Tracking Clinical Results between Visits
6. OP-25 Safe Surgery Checklist Use
7. OP-26 Hospital Outpatient Volume Data on Selected Outpatient Surgical Procedures
8. EDV-1 Emergency Department (ED) Volume

#### **Non-directional measures**

1. MSPB-1/SPP-1 Medicare Spending per Beneficiary (MSPB)
2. OP-9 Mammography Follow-up Rates
3. PAYM-30-AMI Acute Myocardial Infarction (AMI) Payment per Episode of Care
4. PAYM-30-HF Heart Failure (HF) Payment per Episode of Care
5. PAYM-30-PN Pneumonia (PN) Payment per Episode of Care
6. PAYM-90-Total Hip Arthroplasty and Total Knee Arthroplasty (THA/TKA) Hip/Knee Payment per Episode of Care
7. Medicare Hospital Spending by Claim Spending Breakdowns by Claim Type

#### **Measures not required for IQR or OQR**

1. AMI-8a Timing of Receipt of Primary Percutaneous Coronary Intervention (PCI)
2. HF-2 Evaluation of LVS Function
3. PN-6 Initial Antibiotic Selection for Community-Acquired Pneumonia (CAP) in Immunocompetent Patient
4. SCIP-Card-2 Surgery Patients on Beta-Blocker Therapy Prior to Arrival Who received a Beta-Blocker During the Perioperative Period
5. SCIP-Inf-1 Prophylactic Antibiotic Received Within One Hour Prior to Surgical Incision
6. SCIP-Inf-2 Prophylactic Antibiotic Selection for Surgical Patients
7. SCIP-Inf-3 Prophylactic Antibiotics Discontinued Within 24 Hours After Surgery End Time
8. SCIP-Inf-9 Urinary Catheter Removed on Postoperative Day 1 (POD 1) or Postoperative Day 2 (POD 2) with day of surgery being day zero
9. SCIP-VTE-2 Surgery Patients Who Received Appropriate Venous Thromboembolism Prophylaxis Within 24 Hours Prior to Surgery to 24 Hours After Surgery
10. STK-2 Discharged on Antithrombotic Therapy
11. STK-3 Anticoagulation Therapy for Atrial Fibrillation/Flutter
12. STK-5 Antithrombotic Therapy by End of Hospital Day 2
13. STK-10 Assessed for Rehabilitation

14. VTE-4 Venous Thromboembolism Patients Receiving Unfractionated Heparin with Dosages/Platelet Count Monitoring by Protocol or Nomogram
15. OP-31 Cataracts- Improvement in Patient's Visual Function within 90 days Following Cataract Surgery

**Overlapping measures already captured in a composite measure**

1. PSI-6 Iatrogenic Pneumothorax
2. PSI-9 Perioperative Hemorrhage or Hematoma
3. PSI-10 Postoperative Physiologic and Metabolic Derangement Rate
4. PSI-11 Postoperative Respiratory Failure Rate
5. PSI-12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis
6. PSI-14 Postoperative Wound Dehiscence
7. PSI-15 Accidental Puncture or Laceration
8. READM-30-AMI Acute Myocardial Infarction (AMI) 30 Day Readmission Rate
9. READM-30-HF heart Failure (HF) 30 Day Readmission Rate

## **Appendix C: Measures Included in December 2017 Star Rating (N=57) by Group**

### **Mortality**

1. MORT-30-AMI Acute Myocardial Infarction (AMI) 30-Day Mortality Rate
2. MORT-30-CABG Coronary Artery Bypass Graft (CABG) 30-Day Mortality Rate
3. MORT-30-COPD Chronic Obstructive Pulmonary Disease (COPD) 30-Day Mortality Rate
4. MORT-30-HF Heart Failure (HF) 30-Day Mortality Rate
5. MORT-30-PN Pneumonia (PN) 30-Day Mortality Rate
6. MORT-30-STK Acute Ischemic Stroke (STK) 30-Day Mortality Rate
7. PSI-4-SURG-COMP Death Among Surgical Patients with Serious Treatable Complications

### **Safety of Care**

1. HAI-1 Central-Line Associated Bloodstream Infection (CLABSI)
2. HAI-2 Catheter-Associated Urinary Tract Infection (CAUTI)
3. HAI-3 Surgical Site Infection from colon surgery (SSI-colon)
4. HAI-4 Surgical Site Infection from abdominal hysterectomy (SSI-abdominal hysterectomy)
5. HAI-5 MRSA Bacteremia
6. HAI-6 Clostridium Difficile (C. difficile)
7. COMP-HIP-KNEE Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA)
8. PSI-90-Safety Complication/Patient Safety for Selected Indicators (PSI)

### **Readmission**

1. READM-30-CABG Coronary Artery Bypass Graft (CABG) 30-Day Readmission Rate
2. READM-30-COPD Chronic Obstructive Pulmonary Disease (COPD) 30-Day Readmission Rate
3. READM-30-Hip-Knee Hospital-Level 30-Day All-Cause Risk-Standardized Readmission Rate (RSRR) Following Elective Total Hip Arthroplasty (THA)/Total Knee Arthroplasty (TKA)
4. READM-30-PN Pneumonia (PN) 30-Day Readmission Rate
5. READM-30-STK Stroke (STK) 30-Day Readmission Rate
6. READM-30-HOSP-WIDE HWR Hospital-Wide All-Cause Unplanned Readmission
7. EDAC-30-AMI Excess Days in Acute Care (EDAC) after hospitalization for Acute Myocardial Infarction (AMI)
8. EDAC-30-HF Excess Days in Acute Care (EDAC) after hospitalization for Heart Failure (HF)
9. OP-32 Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy

### **Patient Experience**

1. H-CLEAN-HSP Cleanliness of Hospital Environment (Q8)

2. H-COMP-1 Nurse Communication (Q1, Q2, Q3)
3. H-COMP-2 Doctor Communication (Q5, Q6, Q7)
4. H-COMP-3 Responsiveness of Hospital Staff (Q4, Q11)
5. H-COMP-4 Pain Management (Q13, Q14)
6. H-COMP-5 Communication About Medicines (Q16, Q17)
7. H-COMP-6 Discharge Information (Q19, Q20)
8. H-HSP-RATING Overall Rating of Hospital (Q21)
9. H-QUIET-HSP Quietness of Hospital Environment (Q9)
10. H-RECMND Willingness to Recommend Hospital (Q22)
11. H-COMP-7 HCAHPS 3 Item Care Transition Measure (CTM-3)

### **Effectiveness of Care**

1. IMM-2 Influenza Immunization
2. IMM-3/OP-27 Healthcare Personnel Influenza Vaccination
3. OP-4 Aspirin at Arrival
4. OP-22 Emergency Department (ED)-Patient Left Without Being Seen
5. OP-23 Emergency Department (ED)-Head Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Who Received Head CT or MRI Scan Interpretation Within 45 Minutes of Arrival
6. PC-01 Elective Delivery
7. VTE-6 Hospital Acquired Potentially-Preventable Venous Thromboembolism
8. OP-29 Endoscopy/Poly Surveillance-Appropriate Follow-up Interval for Normal Colonoscopy in Average Risk Patients
9. OP-30 Endoscopy/Poly Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use
10. OP-33 External Beam Radiotherapy for Bone Metastases

### **Timeliness of Care**

1. ED-1b Median Time from Emergency Department (ED) Arrival to ED Departure for Admitted ED Patients
2. ED-2b Admit Decision Time to Emergency Department (ED) Departure Time for Admitted Patients
3. OP-3b Median Time to Transfer to Another Facility for Acute Coronary Intervention
4. OP-5 Median Time to Electrocardiography (ECG)
5. OP-18b Median Time from Emergency Department (ED) Arrival to ED Departure for Discharged ED Patients
6. OP-20 Door to Diagnostic Evaluation by a Qualified Medical Professional
7. OP-21 Emergency Department (ED)-Median Time to Pain Management for Long Bone Fracture

## **Efficient Use of Medical Imaging**

1. OP-8 Magnetic Resonance Imaging (MRI) Lumbar Spine for Low Back Pain
2. OP-10 Abdomen Computed Tomography (CT) Use of Contrast Material
3. OP-11 Thorax Computed Tomography (CT) Use of Contrast Material
4. OP-13 Cardiac Imaging for Preoperative Risk Assessment for Non-Cardiac Low-Risk Surgery
5. OP-14 Simultaneous Use of Brain Computed Tomography (CT) and Sinus CT



## Appendix D: Measure Loadings by Group for December 2017

Table D.1. Mortality Measures and Loading Coefficients

Measure Name	Loading Coefficient
<b>MORT-30-AMI</b> Acute Myocardial Infarction (AMI) 30-Day Mortality Rate	0.50
<b>MORT-30-CABG</b> Coronary Artery Bypass Graft (CABG) 30-Day Mortality Rate	0.37
<b>MORT-30-COPD</b> Chronic Obstructive Pulmonary Disease (COPD) 30-Day Mortality Rate	0.65
<b>MORT-30-HF</b> Heart Failure (HF) 30-Day Mortality Rate	0.69
<b>MORT-30-PN</b> Pneumonia (PN) 30-Day Mortality Rate	0.66
<b>MORT-30-STK</b> Acute Ischemic Stroke (STK) 30-Day Mortality Rate	0.51
<b>PSI-4-SURG-COMP</b> Death Among Surgical Patients with Serious Treatable Complications	0.32

Table D.2. Safety of Care Measures and Loading Coefficients

Measure Name	Loading Coefficient
<b>COMP-HIP-KNEE</b> Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA)	0.21
<b>HAI-1</b> Central-Line Associated Bloodstream Infection (CLABSI)	0.02
<b>HAI-2</b> Catheter-Associated Urinary Tract Infection (CAUTI)	0.001
<b>HAI-3</b> Surgical Site Infection from colon surgery (SSI-colon)	0.05
<b>HAI-4</b> Surgical Site Infection from abdominal hysterectomy (SSI-abdominal hysterectomy)	0.05
<b>HAI-5</b> MRSA Bacteremia	0.07
<b>HAI-6</b> Clostridium Difficile (C. difficile)	0.01
<b>PSI-90-Safety</b> Complication/Patient Safety for Selected Indicators (PSI)	0.94

Table D.3. Readmission Measures and Loading Coefficients

Measure Name	Loading Coefficient
<b>READM-30-CABG</b> Coronary Artery Bypass Graft (CABG) 30-Day Readmission Rate	0.25
<b>READM-30-COPD</b> Chronic Obstructive Pulmonary Disease (COPD) 30-Day Readmission Rate	0.54

Measure Name	Loading Coefficient
<b>READM-30-Hip-Knee</b> Hospital-Level 30-Day All-Cause Risk-Standardized Readmission Rate (RSRR) Following Elective Total Hip Arthroplasty (THA)/Total Knee Arthroplasty (TKA)	0.39
<b>READM-30-HOSP-WIDE</b> HWR Hospital-Wide All-Cause Unplanned Readmission	0.97
<b>READM-30-PN</b> Pneumonia (PN) 30-Day Readmission Rate	0.62
<b>READM-30-STK</b> Stroke (STK) 30-Day Readmission Rate	0.51
<b>EDAC-30-AMI</b> Excess Days in Acute Care (EDAC) after hospitalization for Acute Myocardial Infarction (AMI)	0.32
<b>EDAC-30-HF</b> Excess Days in Acute Care (EDAC) after hospitalization for Heart Failure (HF)	0.44
<b>OP-32</b> Facility 7-Day Risk Standardized Hospital Visit Rate after Outpatient Colonoscopy	0.06

**Table D.4. Patient Experience Measures and Loading Coefficients**

Measure Name	Loading Coefficient
<b>H-CLEAN-HSP</b> Cleanliness of Hospital Environment (Q8)	0.76
<b>H-COMP-1</b> Nurse Communication (Q1, Q2, Q3)	0.86
<b>H-COMP-2</b> Doctor Communication (Q5, Q6, Q7)	0.73
<b>H-COMP-3</b> Responsiveness of Hospital Staff (Q4, Q11)	0.82
<b>H-COMP-4</b> Pain Management (Q13, Q14)	0.79
<b>H-COMP-5</b> Communication About Medicines (Q16, Q17)	0.79
<b>H-COMP-6</b> Discharge Information (Q19, Q20)	0.66
<b>H-COMP-7</b> HCAHPS 3 Item Care Transition Measure (CTM-3)	0.90
<b>H-HSP-RATING</b> Overall Rating of Hospital (Q21)	0.91
<b>H-QUIET-HSP</b> Quietness of Hospital Environment (Q9)	0.65
<b>H-RECMND</b> Willingness to Recommend Hospital (Q22)	0.86

**Table D.5. Effectiveness of Care Measures and Loading Coefficients**

Measure Name	Loading Coefficient
<b>IMM-2</b> Influenza Immunization	0.33
<b>IMM-3/OP-27</b> Healthcare Personnel Influenza Vaccination	0.18
<b>OP-4</b> Aspirin at Arrival	0.32
<b>OP-22</b> Emergency Department (ED)-Patient Left Without Being Seen	0.40

Measure Name	Loading Coefficient
<b>OP-23</b> Emergency Department (ED)-Head Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Who Received Head CT or MRI Scan Interpretation Within 45 Minutes of Arrival	0.29
<b>OP-29</b> Endoscopy/Poly Surveillance-Appropriate Follow-up Interval for Normal Colonoscopy in Average Risk Patients	0.52
<b>OP-30</b> Endoscopy/Poly Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use	0.58
<b>PC-01</b> Elective Delivery	0.13
<b>VTE-6</b> Hospital Acquired Potentially-Preventable Venous Thromboembolism	0.19
<b>OP-33</b> External Beam Radiotherapy for Bone Metastases	0.17

**Table D.6. Timeliness of Care Measures and Loading Coefficients**

Measure Name	Loading Coefficient
<b>ED-1b</b> Median Time from Emergency Department (ED) Arrival to ED Departure for Admitted ED Patients	0.84
<b>ED-2b</b> Admit Decision Time to Emergency Department (ED) Departure Time for Admitted Patients	0.78
<b>OP-3b</b> Median Time to Transfer to Another Facility for Acute Coronary Intervention	0.25
<b>OP-5</b> Median Time to Electrocardiography (ECG)	0.22
<b>OP-18b</b> Median Time from Emergency Department (ED) Arrival to ED Departure for Discharged ED Patients	0.79
<b>OP-20</b> Door to Diagnostic Evaluation by a Qualified Medical Professional	0.52
<b>OP-21</b> Emergency Department (ED)-Median Time to Pain Management for Long Bone Fracture	0.40

**Table D.7. Efficient Use of Medical Imaging Measures and Loading Coefficients**

Measure Name	Loading Coefficient
<b>OP-8</b> Magnetic Resonance Imaging (MRI) Lumbar Spine for Low Back Pain	0.04
<b>OP-10</b> Abdomen Computed Tomography (CT) Use of Contrast Material	0.71
<b>OP-11</b> Thorax Computed Tomography (CT) Use of Contrast Material	0.29
<b>OP-13</b> Cardiac Imaging for Preoperative Risk Assessment for Non-Cardiac Low-Risk Surgery	-0.002
<b>OP-14</b> Simultaneous Use of Brain Computed Tomography (CT) and Sinus CT	0.01

## Appendix E: Comparative Analysis of the Updated Star Ratings Methodology

Table E.1 Reclassification of Star Rating with Current and Previous methodologies

Frequency (%)		Methodology Enhancement (December 2017)					Total
		1	2	3	4	5	
Previous Methodology (December 2017)	1	125 (100%)	0	0	0	0	<b>125</b>
	2	135 (19.0%)	575 (81.0%)	0	0	0	<b>710</b>
	3	0	178 (9.1%)	1,187 (60.6%)	594 (30.3%)	0	<b>1,959</b>
	4	0	0	0	561 (68.4%)	259 (31.6%)	<b>820</b>
	5	0	0	0	0	78 (100%)	<b>78</b>
	Total	<b>260</b>	<b>753</b>	<b>1,187</b>	<b>1,155</b>	<b>337</b>	<b>3,692</b>

**Table E.2 Distribution of Group Scores Between Adaptive and Non-adaptive Quadrature**

<b>Mean (Standard Deviation)</b>	<b>Previous Methodology (Non-adaptive Quadrature)</b>	<b>Current Methodology (Adaptive quadrature)</b>
<b>Summary Score</b>	-0.041 (0.504)	-0.016 (0.495)
<b>Group Score</b>	-	-
Mortality	0.002 (0.811)	0.003 (0.812)
Safety	-0.048 (1.046)	-0.044 (1.071)
Readmission	-0.035 (0.978)	-0.039 (1.024)
Patient Experience	-0.122 (1.097)	-0.001 (0.975)
Effectiveness of Care	0.034 (0.727)	0.034 (0.727)
Timeliness of Care	0.005 (0.953)	-0.020 (0.929)
Efficient Use of Medical Imaging	0.022 (0.899)	0.14 (0.902)

## Appendix F: References

1. Lesaffre E, Spiessens B. On the number of quadrature points in a logistic random effects model: an example. *Applied Statistics*. 2001;50:325-335.
2. Capanu M, Gönen M, Begg C. An assessment of estimation methods for generalized linear mixed models with binary outcomes. *Statistics in Medicine*. 2013;32:4550-4566.
3. Pinheiro J, Bates D. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*. 1995;4:12-35.