



September 27, 2017

Kate Goodrich, M.D.  
Director and Chief Medical Officer  
Center for Medicare and Medicaid Services  
7500 Security Boulevard  
Baltimore, MD 21244

Dear Dr. Goodrich:

The Federation of American Hospitals (FAH) is pleased to comment on the proposed methodology changes for the Medicare Hospital Star Ratings program. It is vitally important to hospitals, patients, their families and the overall national work on quality improvement and public reporting that any changes to the display of data by star categories accurately reflect the quality of care provided by hospitals to their patients.

The FAH had serious reservation about the initial Star Ratings methodology and concept. *Hospital Compare* and its measures were never intended to be displayed as composites reflecting a single overall score. Continuing to perpetuate the notion that a single graphic reflects all aspects of hospital care does a disservice to patients, their caregivers and the facilities being measured. The current measures on *Hospital Compare* may not reflect the care of greatest interest to patients, for instance outcomes for emergent surgical care or cancer care follow-up or treatment of a rare disease may be more important than what currently is displayed.

In addition, our members believe it is extremely difficult to achieve a single graphical representation of hospital care using a limited number of variables and statistical constructs that by their nature have limits and may not reflect the overall care delivered in a hospital. However, should CMS continue to try to achieve a graphical representation of quality through star ratings, the FAH strongly suggests CMS further test its methodologies and hold focus groups with hospitals, physicians, patients, families and caregivers to understand how well the statistical information and displays are understood and determined to be useful by all stakeholders.

The FAH also recommends that CMS convene panels of stakeholders to comment on proposed changes to methodologies and to test the understanding of star rating displays to ensure information is conveyed accurately. Any subsequent changes to the star rating methodology should ensure that patients and their families are better able to understand the differences among facilities.

The star rating methodology should be transparent, understandable, and accurately reflect the quality of care provided in the facilities. In other words, the clusters should accurately reflect true differences in care. The current star rating methodology does not do this effectively. Therefore, any changes to the methodology must correct the current flaws and not create additional barriers to patients' understanding of the care provided in hospitals.

Proposed Methodology:

Latent Variable Model: In the proposed changes to the star rating methodology, CMS seeks comment on moving from the current categorization system that uses a winzoration methodology to a new methodology based on  $k$ -means clustering complete convergence stating that  $k$ -means clustering creates a broader distribution of star ratings.

- CMS selects measures from both inpatient and outpatient hospital quality reporting programs and assigns the measures to seven groups: mortality, safety, readmissions, patient experience, timeliness of care, effectiveness of care, and imaging efficiency. Not all measures reported on *Hospital Compare* are used in determining the star ratings. The measure inclusion criteria described in section 5.2 is reasonable. Each group is assigned a weight in the overall star rating.
- A latent variable model (LVM) is used to calculate the measure group score. The LVM approach imputes an unobserved or latent dimension of quality for each hospital that is reflected in the performance measure data. Each of the seven groups of measures has a LVM and a weighted average is taken of those scores to create a summary score for each hospital.
- Star ratings are determined by using a  $k$ -means clustering statistical technique. The  $k$ -means clustering is used to provide assurance that hospital scores within the same star rating are as similar as possible, i.e., scores of hospitals in different star ratings are as different as possible from the star being compared.

Negative Factor Loading: CMS seeks comments on whether or not to include measures with negative factor loading in measure groupings. A negative factor loading determines the extent to which a measure in a group correlates with the overall measure group score. A higher factor loading indicates the measure has a higher correlation with the overall score of the group. Lower factor loadings indicate an inverse relationship between the quality performance score and the group score. The CMS current methodology includes measures with negative factor loadings. The result is that a hospital which performs well on a measure that has a negative factor loading will receive a lower LVM causing the hospital to be penalized, and hospitals that perform poorly on a measure with a negative factor loading will score higher. This type of scoring is problematic in public reporting as it is not clear to the general public nor casual observer of measure scores.

In the proposal for comment, CMS indicates the impact of negative factor loadings on the overall score is relatively modest. Regardless of the impact, the FAH strongly recommends CMS not include measures with negative factor loading in modeling of the star ratings program.

Including measures that react this way in the model is not easily understood by patients, caregivers and the public and does little to inform overall patient care and public reporting.

Measure Group Weightings: The CMS proposal seeks comments on weighting of measure groups. The FAH believes measure group weights should be driven by a clear analytical approach, which includes patient and caregiver preferences. In July 2016, Dr. Francis Vella, Chair of Economics at Georgetown University<sup>1</sup>, conducted a study where the results indicated the weighting of the groups have a significant influence on the star ratings. The FAH encourages CMS to seek patient, family and caregiver input on the weighting of measures with a particular focus on the categories of measures that would be most helpful to them in making decisions about their care. A statistically sound assessment of what patients and caregivers find helpful would better inform the weighting of measure groups.

Public Reporting and *k*-means Clustering: CMS proposes to remove from public reporting and *k*-means clustering those hospitals that do not meet public reporting thresholds. The FAH supports this proposal.

The FAH members encourage CMS to provide mock-ups of the differences in hospital scores and to hold several educational webinars on the proposed changes in the methodology before making the final decision to switch to the new *k*-means clustering methodology. Prior to implementing this change, hospitals should have the opportunity to review their own data run by CMS using this proposed approach. Seeing the exact data that would be displayed will provide hospitals with a better understanding of the impact of the change in display. While hospitals are studying and working to better understand the proposed change in the *k*-means clustering methodology, CMS should continue the suspension the current flawed star ratings.

CMS proposes to run the *k*-means clustering to convergence. The FAH supports this proposal. Running the program to full convergence should indicate clear or obvious demarcation points. If these clusters were not clear, how would CMS decide to assign hospitals to a particular cluster? What is the variation among clusters? Is the variation statistically significant? These are all questions that must be answered prior to implementation of the new methodology.

The methodology, as currently described, does not make clear how extreme outliers were handled and how many extreme outliers were found? The details of the methodology also did not identify which statistical program was used to run the *k*-means cluster analysis, and the options chosen for the statistical procedure. Was it SAS or STATA or another statistical program? Our members with statistical expertise have suggested that the program and the procedure options chosen to run the calculation may make a difference and could influence the final calculation outcome and distribution. CMS should ensure that this information is transparent so that this analysis can be replicated by external stakeholders.

---


<sup>1</sup> Expert analysis of star ratings performed by Dr. Francis Vella, Chair of Economics at Georgetown University published prior to the July 2016 release of star ratings (commissioned by American Hospital Association)  
Attachment 1

The FAH also would appreciate greater detail on how many hospitals were very close to the demarcation lines dividing star levels. In other words, were there five or one hundred hospitals within 0.5 standard deviation of the demarcation line dividing the one-star from the two-star or four-star from three-star categories? The FAH would greatly appreciate CMS releasing the full distribution model.

The FAH appreciates the opportunity to comment on the proposal to move to a new clustering methodology for the Star Ratings. Again, the FAH strongly encourages CMS to continue to suspend the current Star Ratings until a new methodology is in place and hospitals have had the opportunity to fully review their own data and potential display of that data.

Thank you for your consideration.

Sincerely,



Jayne Hart Chambers  
Senior Vice President Quality