March 29, 2019

Kate Goodrich, M.D.
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:

The Federation of American Hospitals (FAH) is the national representative of more than 1,000 investor-owned or managed community hospitals and health systems throughout the United States. Our members include teaching and non-teaching hospitals in urban and rural America, as well as inpatient rehabilitation, psychiatric, long-term acute care, and cancer hospitals. We greatly appreciate the opportunity to comment to the Centers for Medicare & Medicaid Services (CMS) on the February 2019 *Overall Hospital Quality Star Rating on Hospital Compare Public Input Request.*

The FAH appreciates CMS's Public Input Request on potential future methodology changes being considered for the Medicare Hospital Star Ratings program as well as the ongoing efforts to improve the star ratings methodology. 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.

General Comments

The FAH applauds CMS's recognition for the opportunity of a much needed refresh as we continue to hear from our members that, beyond the ongoing methodological issues, a single graphical representation of hospital care using a limited number of measures which are variably reported across hospitals cannot reflect all aspects of hospital care and may mislead the public for whom the tool is intended as a helpful guide.

Moreover, the FAH continues to have reservations about the Star Ratings methodology as the measures it leverages were not developed with the intent to be displayed as part of a composite. In addition, while the statistical methods used to derive the ratings may work well in an exploratory and research capacity, the FAH does not believe application of these methods to

generate a rating to which organizations will be held accountable is prudent. Accountability demands a clear performance target, and not only do the Star Ratings rely on cut-points that are unknown to hospitals in advance they also fluctuate widely. This type of moving target poses challenges to hospitals' understanding of CMS's specific quality performance goals.

The FAH urges CMS to consider alternative ways to construct and present star ratings and to suspend the Star Ratings from the *Hospital Compare* website until concerns with the methodology have been addressed. At a minimum, the methodology should be transparent, understandable, have clear cut-points and targets, and accurately reflect the quality of care provided in the facilities.

To help achieve that goal, the FAH continues to urge CMS to form an additional Technical Expert Panel (TEP) or outside expert group composed of statisticians and biostatisticians who can supplement much needed understanding of the various assumptions and limitations inherent in latent variable modeling (LVM).

Our comments on the specific methodology updates under consideration follow.

Potential Future Methodology Updates

- Measure Grouping: CMS is considering a modified approach to grouping for composing measure groups based on three criteria: an initial clinical grouping, a confirmatory factor analysis, and ongoing active monitoring. The intention is to create a more robust approach that can accommodate changes in measures and hospital performance.

  The FAH supports an approach that ensures periodic re-evaluation of the measure groups, to properly account for the measures being added to and removed from the Star Ratings measure set and to ensure that measure loadings are balanced and positive. As we noted in comments to the Hospital Inpatient Prospective Payment System FY 2019 proposed rule, it is important to consider the impact on Star Ratings when CMS proposes retiring measures from hospital quality reporting programs. While the periodic confirmatory factor analysis would be crucial to provide more empirically sound and consistent measure groups CMS should also consider how measure would cause disruptions to the Star Ratings if removed and provide information on the impact to the ratings if such changes are being considered for public comment.

- Potential Regrouping of Patient Safety Measures: CMS is considering two options for regrouping the patient safety measures when calculating the Star Ratings. Option 1 involves the retention of the PSI-90 composite. Option 2 involves breaking down the PSI-90 into its component measures. CMS requests comment on the suitability of the current or either of the two alternative grouping options.

  FAH believes that neither grouping option (current or alternatives) is most suitable because they all continue to rely on the PSI-90 measure or its components. FAH continues to urge CMS to consider the removal of these measures from quality programs given ongoing issues with the reliability and validity of the PSI-90

composite and its underlying components. Although FAH recognizes that it is important to include as many measures as possible in the Star Ratings, there is no benefit to including measures that do not result in an appropriate assessment of hospital performance.

Beyond this, the FAH notes that in both options presented there are issues with achieving balanced loadings and as such neither option is ideal. The subdivision of the PSI-90 into its component measures would at least increase the level of specificity fed into the model as surgical and medical adverse events require different approaches for improvement. In addition, the direct connection between specific measures and overall ratings allow hospitals to aim for more targeted performance improvement activities with physicians. However, a concern with using the component measures is that the contribution of hospital-associated infection (HAI) measures might be suppressed. The HAI and other safety events are such low frequency events that there is little predictive value from quarter to quarter. **The FAH requests that CMS reveal how it would account for these data if PSI-90 is broken down into its component measures, and how CMS would contend with the low predictive values of safety event measures.**

- Incorporating Precision of Measures: CMS is considering changing the weighting options that account for differences in measure score precision across hospitals. The current methodology uses denominator weighting. This methodology has contributed to imbalances in the loadings, causing some measures to be more heavily weighted than others to the detriment of the consideration of hospital performance on the overall star rating.

  If CMS continues with the current methodology, applying log transformations in the denominators provides a more equitable distribution of loadings. However, this approach, while methodologically preferable in this context, is not intuitive and will be difficult to explain to stakeholders. **The FAH cautions against using a mixed weighting methodology across the different measures and urges CMS and its contractor to evaluate thoroughly the impact of any change in methodology it is proposing and to share that information with stakeholders for review and comment prior to implementation.**

- Reporting Schedule and Period-to-Period: **The FAH supports shifting to annual reporting ratings from the current biannual schedule.** This would provide increased stability to the ratings and would be consistent with the schedule of annual updates that are reported for most outcome measures. While one drawback to an annual assessment is that it limits the visibility of changes that may be improving or worsening scores when measure updates that do not fall within the yearly update take place. However, this lag would be no longer than 9 months in the worst case.

The FAH does not support the use of weighted or time-average previous period data in calculating the Star Ratings. A hospital's past performance may not be the best predictor of current or future performance hence use of older information may well result in ratings that are not relevant to consumers who may rely on the star rating to choose a hospital for their care. Indeed, data lags for some measures already limit how current a reflection the Star Ratings provide.

The FAH supports exploration of an alternative way of reducing period-to-period shifts through the use of three-star or partial-star categories rather than five-star. Three star categories would provide patients with information on outliers which is helpful in guiding consumer choice while likely introducing improved consistency from period-to-period. Partial star ratings might provide more clarity if implemented correctly. The FAH supports empiric evaluation and consumer testing of such an approach.

- Peer Grouping – CMS currently publishes ratings across all hospitals regardless of hospital characteristics such as range of services provided or populations served leading to concern that this does not constitute an apples-to-apples comparison. CMS seeks feedback on peer grouping in order to allow comparisons across hospitals that share key attributes understood to influence the rating. **The FAH supports peer grouping across dual eligibility status as a first step towards improved risk adjustment. However, risk adjustment itself is necessary and CMS should continue to work toward implementing that.** A fully defined socioeconomic status risk adjustment method is preferred.

  Although FAH supports peer grouping to allow comparisons across hospitals with shared characteristics, FAH is concerned that expanding the use of peer grouping to include multiple levels of stratification in addition to proportion of population with dual eligibility would likely complicate interpretation of the Star Ratings for consumers. **FAH urges CMS to test any potential stratified comparisons of star ratings among hospitals, physicians, patients, families, and caregivers and seek their feedback prior to any implementation.**

Potential Long-Term Methodology Changes

- Replacing LVM with an explicit approach to group score calculation: **The FAH strongly supports replacing LVM with a simpler, more explicit approach to group score calculations that yields a more intuitive and predictable approach to describing hospital quality performance.** CMS currently uses a latent variable modeling (LVM) and k-means clustering to compress 57 measures into 5-star ratings. The complexity of the LVM model and combined k-means clustering, while methodologically elegant, leads to unpredictability of the group scores from reporting period to reporting period. This makes it difficult for hospitals to understand the current factors contributing to those scores and take appropriate actions, limiting its utility as an approach for scoring measures and measure sets intended for performance rating.

These methods have also resulted in misleading ratings of hospital quality which does a disservice to patients, their caregivers and the facilities being measured. The rating should be intuitive with directionality of performance measure scores for it to be actionable for hospitals. The relationship between a final score and the measures that are its building blocks should not be inscrutable nor should future performance be unpredictable for the organization being measured. The advantages of a more explicit approach to Star Ratings include predictability that allows hospitals to estimate their future performance.

An explicit approach would be an improvement over the LVM, and FAH recognizes that this will require policy decisions when applying weights and including measures. However, as far as specific approaches, the FAH cautions against a simple averaging approach, in particular, if no confidence intervals are used.

- Clustering alternatives: The FAH is not opposed to the use of k-means clustering as a method for stratifying. However, use of k-means clustering for assigning star ratings faces the considerable issues that CMS has identified from stakeholder feedback: 1) the inability to predict cut points is severely problematic for hospitals in a rating measure, and 2) clustering real-life, not normally distributed, messy data tends to lead to suboptimal clusters. While imperfect or non-intuitive methods of clustering work well for segmenting populations, they are flawed when it comes to establishing ratings that are publicly reported or tied to payment. For these purposes, a hospital should have a precise performance standard to target and it needs to be able to estimate its performance against that standard. If there is within-hospital consistency across quality measures over time, there should also be within-hospital consistency over time in its star rating score.

  Clusters should accurately reflect true differences in care. **Regardless of the graphical representation, FAH urges CMS to test any changes by holding 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.**

- Incorporation of Improvement: **The FAH does not support incorporation of hospital-specific improvement into Star Ratings.** If a hospital's star rating changes from a previous period, it will either be rewarded or disadvantaged by the new rating. There is no need to explicitly include the change in calculating the rating. Public reporting of an indicator of directionality of change would contribute to consumer confusion and may not be meaningful information. As noted above, a change in one year may not be predictive of current or future performance.

The FAH appreciates the opportunity to comment on the options and proposals to move to an improved Overall Hospital Quality Star Rating. If you have any questions regarding our comments, please do not hesitate to contact me or Claudia Salzberg of the FAH staff at (202)624-1500.

Sincerely,

Chip Kahn
President and CEO

Cc: Michelle Schreiber, M.D., CMS
Reena Duseja, M.D., CMS