Association of
American Medical Colleges
655 K Street, NW, Suite 100, Washington, DC 20001-2399
T 202 828 0400
aamc.org

Via Email (cmsstarratings@yale.edu)

March 28, 2019

Michelle Schreiber, M.D.
Director, Quality Measurement and Value-Based Incentives Group
Center for Clinical Standard and Quality
Centers for Medicare & Medicaid Services
U.S. Department of Health and Human Services
7500 Security Boulevard
Baltimore, MD 21244

RE: Overall Hospital Quality Star Rating on* Hospital Compare *Public Input Request

Dear Dr. Schreiber:

The Association of American Medical Colleges (AAMC) appreciates the opportunity to comment on the public input request to provide feedback on potential updates and future consideration for the methodology of the Overall Hospital Quality Star Rating on *Hospital Compare*, issued by the Centers for Medicare & Medicaid Services (CMS).

The AAMC is a not-for-profit association dedicated to transforming health care through innovative medical education, cutting-edge patient care, and groundbreaking medical research. Its members are all 154 accredited U.S. and 17 accredited Canadian medical schools; nearly 400 major teaching hospitals and health systems, including 51 Department of Veterans Affairs medical centers; and more than 80 academic societies. Through these institutions and organizations, the AAMC serves the leaders of America's medical schools and teaching hospitals and their more than 173,000 full-time faculty members, 89,000 medical students, 129,000 resident physicians, and more than 60,000 graduate students and postdoctoral researchers in the biomedical sciences. Together, these institutions and individuals are the American academic medicine community.

The AAMC appreciates the CMS dedication of future time and work on improving Star Ratings. We remain very concerned with the flawed methodology used to determine the Ratings posted on *Hospital Compare* and believe them to be both inaccurate and misleading to patients and consumers seeking hospital care. We urge CMS to continue to engage stakeholders throughout the Ratings improvement process.

Summary of Key AAMC Recommendations

The following items are the AAMC's key recommendations on methodologic improvements:

- ***Suspend the Star Ratings***: CMS should remove the publication of the Star Ratings from the *Hospital Compare* website until CMS is able to address significant concerns with the methodology.

- ***Improve Underlying Measures***: CMS should improve existing measures in use in the hospital quality reporting and performance programs, including the incorporation of sociodemographic factors in measure-level risk adjustment. CMS should remove PSI-90 from the Star Ratings.
- ***Overall Composite Ratings Add to Confusion About Hospital Confusion***: A rating that combines all of the multiple dimensional aspects into a summary score may not provide a patient or consumers with the information that is truly important for an individual's situation. The AAMC urges CMS to explore the template matching, or other approaches that directly compare patient groups, as a possible alternative model to use for rating hospitals.
- ***Measure Groupings***: CMS should undertake further analysis on how to improve measure groupings before implementing any changes.
- ***Regrouping Measures***: The AAMC does not support the proposed alternative grouping for the Safety of Care group or the use of PSI components in lieu of the PSI-90 composite measure. CMS should consider simpler alternative approaches before implementing any regrouping of measures.
- ***Incorporating Precision of Measures***: CMS should continue to analyze and share more information on potential approaches to improving the incorporation of measure precision in the ratings and be transparent in advance of implementing changes.
- ***Period-to-Period Star Rating Shifts***: The AAMC supports the proposal to move to an annual update cycle, until there are further improvements in the reliability and stability of the methodology, and further exploration of "partial-star" Overall Hospital Ratings. We believe the drawbacks outweigh the possible benefits of incorporating weighted averages to address period-to-period shifts.
- ***Peer Grouping***: The AAMC remains supportive of peer grouping and believes stratified comparisons are useful to patients and consumers to best understand the different types of hospitals available to them, especially as a short-term solution to the broader need to develop more rigorous risk adjustment at the measure-level. We urge CMS to ensure that stratified comparisons of hospital performance are clear when published on *Hospital Compare*.
- ***Explicit Approach***: The AAMC supports additional work around the development of an explicit approach to replace the Latent Variable Modeling.
- ***Clustering Alternative:*** CMS should consider alternatives to the current k-means clustering, with a focus on an approach that allows for predictable, fixed targets.
- ***User-Customized Star Rating***: The AAMC is supportive of exploring user-customization to the Star Ratings, but believes additional investigation and analysis is needed to better understand patient and consumer interest and to ensure that customized ratings are reliable and valid.

GENERAL COMMENTS

Suspend the Star Ratings Until Flaws are Addressed

The AAMC calls on the Administration to remove the publication of the Star Ratings from the *Hospital Compare* website until CMS is able to address significant concerns with the methodology. We request that prior to releasing Star Ratings, CMS take sufficient time to examine the feedback provided and make modifications to the methodology to ensure that the Ratings are accurate. We remain extremely concerned about potential consequences for patients that could result

from an overly simplistic picture of hospital quality with a single overall rating. It is imperative that CMS contract with independent outside experts to review the methodology and verify its accuracy before public implementation.

The AAMC also strongly recommends that CMS continue ongoing review for areas of improvement in future releases of the Ratings and convene stakeholders regularly to review the appropriateness of the current methodology.

Improve the Underlying Quality Measures

An overall quality rating based upon individual quality measures can only ever be successful if the underlying measures themselves are reliable, valid, and incorporate appropriate and robust risk adjustment to accurately account for the differences in clinical and social risk of patients that a hospital serves. **The AAMC urges CMS to improve upon existing measures in its hospital quality reporting and performance programs while also undertaking efforts to update and improve the Star Rating methodology.**

Incorporate Sociodemographic Status (SDS) Factors into Measure-Level Risk Adjustment
Approximately two-thirds of a hospital's Star Rating is based on its readmissions, mortality, and patient experience performance. There is significant peer-reviewed literature[1] demonstrating that hospital performance on these outcomes can be affected by factors outside the control of the hospital (e.g., housing, food insecurity, social support, and transportation). Furthermore, Congress recognized that hospitals that disproportionately care for vulnerable patient populations, who are at a higher risk of readmissions, are disadvantaged when these factors are not considered in the payment scoring methodology and mandated that CMS adjust hospital readmission penalties for the proportion of dually eligible patients under the Hospital Readmission Reduction Program. CMS has implemented this adjustment through stratifying penalties by the proportion of Medicare and Medicaid dual-eligible patients the hospital serves. This stratification is only the first step toward accurate risk adjustment for patients with social and economic challenges. CMS must go beyond adjusting only payments to also adjusting the underlying measures in order to make accurate quality comparisons.

One promising avenue for incorporating SDS factors into measure-level risk adjustment is the National Quality Forum (NQF)'s NQP Social Determinants of Health Data Integration Project which ensures that measure developers are improving measures currently in use by incorporating critical SDS data elements into measure risk adjustment when possible. The AAMC urges CMS to work with NQF on this effort.

Remove PSI-90 from Star Ratings
The AAMC has numerous concerns with the PSI-90 composite measure. Some of the components of the measure focus on surgical care, which disadvantages teaching institutions that tend to have a larger volume of surgical cases than do other hospitals. Further, the PSI-90 tends to penalize hospitals that have large volumes of surgeries, even where the probability of an adverse event is the same as a low-volume hospital. Additionally, some components of the measure are susceptible to surveillance bias and therefore institutions that are more diligent about reporting safety events are

[1] *See* National Academies of Sciences, Engineering, and Medicine. 2016-2017. *Report Series: Accounting for Social Risk Factors in Medicare Payment*. Washington, DC: The National Academies Press. Details here: http://www.nationalacademies.org/hmd/Activities/Quality/Accounting-SES-in-Medicare-Payment-Programs.aspx

penalized[2, 3] For example, teaching institutions tend to have robust infection control programs, which focus on identifying and reporting patient safety events. Finally, the measure is based on administrative claims data so cannot capture the full scope of patient-level risk factors.[4,5,6] While the modified composite may be an improvement over the previous version, many of the issues previously cited in comments to the Agency continue to apply, and because of this, **CMS should remove the PSI-90 measure from the Star Ratings methodology**.

An Overall Hospital Compare *Composite Rating Adds to Confusion About Hospital Quality*

The AAMC strongly supports making quality data available in an easy to understand format for patients and the public. While we support efforts for greater transparency, we believe that this information must be displayed in a meaningful fashion. **A single composite rating that combines disparate quality measures, particularly those that lack clinical nuance, oversimplifies the complex factors that must be taken in account when assessing the care quality. The hospital star ratings are not a useful metric of overall quality of a hospital but a metric of a few discreet processes of questionable representation of overall quality and most importantly outcomes.** This is particularly true for the nation's teaching hospitals that typically care for sicker and more vulnerable patients in a diverse and complex environment.

Rather than using a single composite score methodology, the AAMC recommends the development of Ratings for subsets of measures, which should ultimately be more meaningful and actionable for both patients and consumers, but also for the hospital's quality improvement efforts. The measures on *Hospital Compare* cover a wide variety of conditions and procedures for the inpatient, outpatient, and emergency department settings yet under the current methodology only a handful of scores ultimately determine a hospital's overall quality rating and compares hospitals regardless of the number of measures the hospital is scored on or services the hospital offers. **A rating that combines all of the multiple dimensional aspects into a summary score may not provide a patient or consumers with the information that is truly important for an individual's situation.** Even worse, the current system does not shine light on the differences between hospitals compared or disclose the areas where a given hospital might not provide a given service or may lack a measure score. Patients may choose a hospital for a particular condition or location at one time, and may make a different choice at another time and should have better access to quality information to inform those choices. We are concerned that patients lack the multifaceted information they need to aid them

---

[2] Koenig, Lane et al. *Complication Rates, Hospital Size, and Bias in the CMS Hospital-Acquired Condition Reduction Program*. American Journal of Medical Quality. December 19, 2016. Retrieved from: https://journals.sagepub.com/doi/abs/10.1177/1062860616681840.

[3] Blay Jr., Eddie et al. *Evaluating the Impact of Venous Thromboembolism Outcome Measure on the PSI 90 Composite Quality Metric*. The Joint Commission Journal on Quality and Patient Safety. March 2019. Retrieve from: https://www.jointcommissionjournal.com/article/S1553-7250(18)30220-4/pdf

[4] "MedPAC Comments on FY 2014 IPPS Proposed Rule." June 25, 2013. Retrieved from: http://www.medpac.gov/documents/comment-letters/medpac's-comment-on-cms's-acute-and-long-term-care-hospitals-proposed-rule.pdf?sfvrsn=0

[5] Rajaram, Ravi et al. *Concerns About Using the Patient Safety Indicator-90 Composite in Pay-for-Performance Programs*. JAMA. Vol 313, No. 9. March 3, 2015. Retrieved from: http://jama.jamanetwork.com/article.aspx?articleid=2109967

[6] *Medicare's Hospital-Acquired Condition Reduction Program*. Health Affairs: Health Policy Briefs. August 6, 2015. Retrieved from http://www.healthaffairs.org/healthpolicybriefs/brief.php?brief_id=142

in their healthcare choices. Distilling a large amount of information into one overall rating is not useful.

The Potential of a Template Matching Model as an Alternative Approach
The current method of measuring hospital outcomes on *Hospital Compare* primarily focuses on an indirect standardization, where a hospital's own case mix is used for comparing performance. This approach compares hospitals that maintain important differences in patient populations served (both in complexity and in social risk factors).

An alternative approach could combine the benefits of indirect standardization with the appropriateness of direct standardization, which seeks to compare hospitals relative to an external reference population. This may be more meaningful for patients in that such a method would be more reliable for defining how well the hospital has done with other patients who have the condition for which they are seeking care.[7], [8] A mixed approach, known as a "hospital-specific template matching method", recently developed by researchers, seeks "to better implement indirect standardization analyses for improving a hospital's quality of care specifically tailored to the index hospital's most relevant patients – the patients they see."[9] Under this approach, they have found that the method "combines the fairness of comparison from direct standardization with the specific institutional relevance of indirect standardization." Considering that the Hospital Compare Overall Quality Star Rating is meant to assist patients and consumers choose hospitals based upon quality information *and* help guide hospitals in their quality improvement activities, the template matching model may be a valid alternative worthy of full consideration. **The AAMC urges CMS to explore the template matching, or other approaches that directly compare patient groups, as a possible alternative model to use for rating hospitals.**

POTENTIAL FUTURE METHODOLOGY UPDATES

Measure Grouping

CMS seeks feedback on using an explicit three-step approach to define measure groups that might be reasonable to ensure that measure groups are both clinically and empirically rational. CMS is proposing a new approach to measure grouping based upon three criteria: initial clinical grouping, confirmatory factor analysis, and ongoing active monitoring. The reasoning behind this proposal is that in part the Agency has begun to retire measures from the Inpatient Quality Reporting Program and other hospital reporting and performance programs as part of its broader Meaningful Measures Framework, and recognizes that changes to the measures reported on *Hospital Compare* could have an impact on the current measures groups utilized in the Star Ratings methodology. **The AAMC believes that the three-step approach to define measure groups is reasonable, but CMS should undertake further analysis on how to improve measure groupings before implementing any changes.** Our concerns are discussed in further detail in the following section in regard to the impact such an approach would have on the Safety of Care measure group.

---

[7] *See* Silber, JH et al. *A hospital-specific template for benchmarking its cost and quality.* Health Services Research. October 2014. Retrieved from: https://www.ncbi.nlm.nih.gov/pubmed/25201167.
[8] *See also* Silber JH et al., *Comparison of the Value of Nursing Work Environments in Hospitals Across Different Levels of Patient Risk.* JAMA Surgery. June 2016. Retrieved from: https://www.ncbi.nlm.nih.gov/pubmed/26791112.
[9] *See* Silber et al., *A hospital-specific template for benchmarking its cost and quality,* 1477

Additionally, CMS also asks for feedback on using balance and consistency of loadings as a factor in evaluating measure grouping. **As a guiding principle, the AAMC agrees that balance and consistency of measure loadings are important additional factors for evaluating measure groupings.**

Regrouping Measures

Following upon the discussion of a three-step approach to measure grouping, CMS notes that the Safety of Care group might need reconsideration as the application of the criteria suggest that the underlying latent variable may be weaker compared to other measure groups. CMS seeks feedback on its hypothesis that dividing the Safety of Care group into two separate clinical safety groups, one for *surgical* safety and the other for *non-surgical* (or medical) safety, might lead to greater stability with a stronger underlying latent variable.

The AAMC understands the Agency's desire to improve the Safety of Care group, but questions the concept of splitting the group into two and specifically the proposed break between safety of surgical and medical services. For one, it is challenging to split the health-care acquired infections between surgical and medical services. For example, a central line-associated bloodstream infection (CLABSI) or a catheter-associated urinary tract infection may ultimately be the result of a central-line or a catheter inserted during surgery, but could also follow medical cases. Additionally, is there potential for confusion on the part of patients and consumers if measures are split between surgical and medical safety and rated separately? Has CMS run analysis to determine whether the break will increase the number of hospitals without measure group scores? **As proposed, the AAMC does not support the potential alternative groupings for the Safety of Care measures and does not believe it to be suitable for the Overall Hospital Quality Star Rating.** We also encourage CMS to share future analysis publicly for additional input.

Additionally, as part of the proposal to split the Safety of Care measure group into two separate groups, CMS seeks feedback on also splitting the PSI-90 composite measure and measuring its components. This would allow CMS to account for the surgical components under the new surgical safety group (eight of ten components) and the non-surgical components under the new medical safety group, rather than assigning the PSI-90 composite entirely to the surgical safety group. The AAMC has several concerns about this proposal. First, splitting PSI-90 into components further complicates the split into medical and surgical groupings, since many of the components are not exclusively medical or surgical. For example, pressure ulcers (designated as medical) could result from the required rest following a complicated surgery or the rate of postoperative respiratory failure (designated as surgical) is influenced more by a patient's co-morbidities than the surgery itself. Second, breaking PSI-90 into components may decrease the reliability of the measures by assigning scores to very rare events. We encourage CMS to provide information about the statistical significance of the measured rates of individual PSI measures at the hospital level. Further, the denominators may vary drastically across hospitals and could exacerbate the biases seen in the PSI-90 composite.

Even at the national level, individual components saw huge swings in weight across time. In Table 8CMS showed that not only did four components change by more than 0.20 across periods, but they also caused changes in other measures in the same group, such as hip/knee complications, which

changed by over 0.50 across periods. It is noteworthy that these changes occurred despite no updates to the hip/knee complications measure itself, such as between July 2017 and December 2017. Thus, we are concerned that breaking down PSI-90 into individual components could further destabilize the Safety of Care group. **The AAMC does not support the use of the PSI components in lieu of the PSI-90 composite measure and would encourage caution in using measures that bring reliability into question.** The AAMC continues to believe that the PSI-90 composite measure should be removed from the Star Ratings.

An alternative CMS would be to implement a simpler approach that focuses on consistent and balanced measure loadings. Such a model would increase interpretability and add needed balance across the measures of a unified Safety of Care measure group, and remove the need to split the group into two. **The AAMC urges CMS to consider simpler alternative approaches before implementing any regrouping of measures.**

Incorporating Precision of Measures

Currently, CMS uses a denominator weighting to account for differences in measure score precision. Further analysis of this approach has revealed that in addition to reflecting sample size differences, denominator weighting may also contribute to the imbalance of measure loadings and worsen model fit, but that the cause of this effect is unknown. CMS has considered three alternative weighting options to account for precision of the measure: (1) Confidence interval weighting; (2) Log (denominator) weighting for non-volume denominators, otherwise use of denominator weights; and (3) No weighting (equal weighting). CMS notes that none of options is without disadvantages (primarily expected shifts in ratings or lack of intuitive support), but believes that incorporating measure precision into the ratings is conceptually important. We note that the alternative approaches often demonstrated large, unexplained changes in measure loadings over time, such as for the hip/knee complication rate in the confidence interval weighting. We are concerned that CMS does not fully understand the reasons for differences across different denominator weighting, and caution against any action before further analysis. **The AAMC agrees that measure precision is critical to the ratings, but insufficient data and specific details are available to assess the options. CMS should continue to analyze and understand approaches to improving the incorporation of measure precision in the ratings, and be transparent in advance of implementing any changes.**

Period-to-Period Star Rating Shifts

In response to the substantial shifts in ratings observed in the unpublished July 2018 release, CMS has undertaken analysis of options to stabilize period-to-period shifts in the ratings. From this, CMS seeks feedback on the following potential improvements: (1) use of a weighted average summary score, (2) use of "partial" Star Ratings, and (3) moving to an annual refresh schedule. Shifts in ratings observed from measurement period to measurement period cause the ratings to appear random, and thus are difficult for hospitals to use for performance improvement activities. The AAMC remains concerned about these shifts, and our comments to each of CMS's potential improvements are below.

Incorporating data from an older period, especially at higher weights, would drastically reduce shifts of two or more stars. The drawback to such an approach is that it would limit the timeliness and

currency of data available to patients and consumers and delay a hospital's realization of improvement in the ratings relative to improvement observed on the underlying measures. **The AAMC believes the drawbacks outweigh the possible benefits, and does not support the incorporation of a weighted average to address period-to-period shifts.**

An alternative to use of a weighted average is the incorporation of "partial" or "half" Star Ratings, such as 2.5 stars or 3.5 stars, as this would reduce the "cliffs" between hospital categories (i.e., the actual difference in scores between a "high" 2-star and a "low" 3-star hospital) and provide greater clarity to patients and consumers on a hospital's relative performance. Additionally, the Agency's Home Health Compare Star Ratings for home health agencies and the Medicare Advantage Plan Quality Star Ratings programs utilize half-star ratings, suggesting that similar methodologies for use of half-stars might be easier to implement. **The AAMC supports further exploration of "partial" Star Ratings, as we agree that it may be an appropriate alternative option to reduce period-to-period shifts.**

Finally, CMS seeks comment on whether it should move to an annual update cycle for the Overall Hospital Quality Star Rating, essentially tying the timing of the ratings cycle to measures that are refreshed annually (which include most of the underlying outcomes measures: PSI-90, hip/knee complications, EDAC measures, readmissions measures, and mortality measures). Stakeholders have previously expressed concern that the current biannual ratings update is not aligned with annual measure refreshes, and may result in changes in rating for hospitals near cutoffs due to sensitivity to modest changes to measures outside the major annual refresh schedule. Given the current issues and concerns with the methodology, moving to an annual refresh schedule would smooth period-to-period shifts and provide greater predictability in the release schedule. **The AAMC supports the proposed move to an annual schedule for the Overall Quality Star Ratings until there are further improvements in the reliability and stability of the methodology.**

Peer Grouping

CMS seeks feedback on the value of and ways it should calculate Overall Quality Star Ratings among peer groups, in an effort to present the ratings results based on hospitals that "look like them." As currently implemented, CMS compares all hospitals that meet the minimum measure requirements (nine measure scores, across a minimum of group measure groups, with at least one measure group related to outcomes) regardless of differences in hospital characteristics, such as teaching or safety-net status, number of beds, or range of services provided. Teaching hospitals perform a wide array of complicated and common procedures, pioneer new treatments, and care for broader socio-demographic patient populations that may have limited access to care. Yet under the current Star Ratings program, they are compared directly to hospitals with homogenous patient populations and hospitals that do not perform enough procedures to be measured on a majority of the individuals included in the methodology. This had led to observations that the ratings disadvantage large teaching hospitals.

The AAMC has consistently supported peer grouping as a way to stratify the ratings by hospital type or characteristic, and has previously recommended that CMS explore measure performance within specific hospital peer cohorts so that hospitals with similar characteristics and risk profiles are compared to each other. **The AAMC supports peer grouping and believes stratified comparisons are useful to hospital stakeholders for quality improvement activities and also to help patients**

and consumers best understand differences among the various types of hospitals available to them. CMS's Nursing Home Compare Star Ratings and Star Ratings for Medicare Advantage and Part D plans, in addition to the Veteran's Affairs Strategic Analytics for Improvement and Learning (SAIL) Hospital Star Ratings, account for differences in cohort being compared. The AAMC believes that CMS should look to these programs to inform peer grouping in the Overall Hospital Quality Star Ratings. The AAMC asks that CMS consider multiple stratification approaches and share analysis or data simulation of different approaches, to help inform stakeholder feedback.

Variables

CMS seeks feedback on the variety of variables it could use for peer grouping (proportion of dual-eligible patients, number of measures reporting, teaching status, number of beds, specialty, critical access hospital, for example) and which of those would be most useful. The AAMC understands that each variable may have advantages and disadvantages, and that no one variable for peer grouping will address the lack of adequate risk adjustment to account for SDS factors. We recommend that, until a more refined methodology is available, CMS stratify by either social risk, using proportion of dual-eligible patients similar to the peer grouping implemented in the Hospital Readmission Reduction Program for symmetry, or by hospital size/full service status, to ensure patients are able to compare hospitals that are able to fully meet their care needs. **Regardless, the AAMC asks CMS to implement peer grouping as a short-term solution while it addresses the broader need to develop more rigorous risk adjustment at the measure-level. CMS should conduct a thorough analysis of the extensive data it has available to determine the most appropriate peer groups.**

Presentation of Peer Grouping

CMS notes that there is disagreement among stakeholders on how peer grouping would be presented on the Hospital Compare website. In particular, some stakeholders believe that peer grouped results would be confusing and unhelpful to patients and consumers and thus peer grouped results should be presented as supplemental information to the unstratified Overall Hospital Quality Star Rating. The AAMC disagrees that such information is confusing or unhelpful. Instead, we believe patients should be able to discern the range of services available at any particular hospital, and whether that hospital has reported measures of importance to the patient. As currently presented, a patient might not be able to distinguish whether the "top rated" hospital in the patient's region is a community hospital that may be unable to care for patients with more complex conditions. Presenting the peer grouped ratings within the web-based tool when a patient searches for hospitals will assist that patient in better understanding the options available. **The AAMC urges CMS to ensure that stratified comparisons of hospital performance are clear when published on *Hospital Compare*.**

POTENTIAL LONG-TERM METHODOLOGY CHANGES

Explicit Approach

CMS is considering replacing the latent variable modeling (LVM) with a less complex or more explicit approach. The LVM was chosen in part to reduce arbitrariness, but as a disadvantage the LVM introduces inherent uncertainty into the ratings at hospitals because the measure loadings are unknown until data is refreshed and may change over time. This leads to little transparency or predictable advance notice for hospitals in how changes in individual measure scores may impact hospital Star Ratings.

CMS describes different ways it could implement an explicit approach. One discussed in the public input request is to keep the current methodology except that instead of the LVM to determine measure loadings, CMS would assign weights to each measure within a group. While this would provide greater transparency and predictability to hospitals, it would require broad stakeholder agreement on which measures to weight more heavily, or whether to weight all measures equally. Furthermore, CMS believes such consensus might be difficult to achieve, especially over time as measures may change. The LVM approach, while not transparent or predictable, may be more feasible to maintain over time as it responds to the data based on the correlations between measures each refresh to calculate measure loadings. In response to this public input request, we anticipate others will propose specific, technical alternative approaches to the LVM. The AAMC asks that CMS share these proposals, and produce a comparative analysis to which stakeholders can respond to.

The AAMC agrees that an explicit approach is likely to be easier to understand for hospitals and patients alike, introduce predictability and transparency to the ratings, and allow for a greater balance and consistency of measure weights. We acknowledge that gaining consensus on measure contribution weights would likely be difficult, but that process may ultimately result in greater stakeholder "buy-in" on the Overall Hospital Quality Star Ratings. One way to operationalize such consensus is to convene a single, inclusive advisory group, rather than separate work groups separating stakeholders, whose deliberations are open to the public and whose recommendations to CMS are subject to public comment. Such a group could be formed initially to provide feedback on the development of an explicit approach and retained for annual reviews for maintenance of the approach. **Overall, the AAMC supports additional work around the development of an explicit approach for CMS to consider further. In particular, the AAMC urges CMS to consider template matching, and other models that allow for a direct comparison of hospitals, for additional exploration** (as described earlier in these comments)**.**

Clustering Alternative

Currently, CMS uses a k-means clustering method to assign hospitals to a discrete Star Rating category. This approach was originally used to avoid arbitrary cut points, accommodate changes to the underlying distribution of scores, and to provide a comparative assessment for patients. However, such an approach also limits hospitals' ability to predict cut points for future releases and seems arbitrary for hospitals with borderline scores. CMS seeks feedback on whether it should consider potential alternatives to k-means clustering and what sorts of changes.

In many other areas of performance measures, those being scored have knowledge of a fixed target one must achieve in order to meet the "grade" one desires. The AAMC believes that a "line of sight" between a hospital's performance and its star rating is critical to the future utility of the Overall Hospital Quality Star Ratings continuing to motivate quality improvement. Explicit predictable scoring targets are key drivers for hospitals to invest in meaningful improvement activities. To that end, **we believe that CMS should consider alternatives to the current k-means clustering, with a focus on an approach that allows for predictable, fixed targets.**

User-Customized Star Rating

Currently the Star Ratings are based upon fixed measure group weights, representing a generalized vision of aspects of quality that are important to measure, while allowing hospitals to be compared against each other under a common rubric. These group weights, however, may not capture priorities, preferences, or values of an individual patient or consumer. CMS seeks feedback on whether to further explore the introduction of a user-customization tool to the Overall Hospital Quality Star Rating, and on how to build and implement such a tool.

The AAMC agrees that the Overall Hospital Quality Star Ratings must be meaningful to patients and consumers. We agree that a customizable ratings tool conceptually might create greater alignment with the consumer focus of the Ratings. CMS discusses a single measure group weight customization concept as one way to generate user-customized Star Ratings.

The measure group weight customization concept described appears to follow the work by the RAND Corporation with its Personalized Hospital Performance Card.[10] Under RAND's concept, a user can see the difference among hospitals' Star Ratings under the prescribed measure group weightings. The user also can manipulate those group weightings. For example, a user could determine she only cares about mortality, and re-weights mortality 100% of the Rating, and compare that result to CMS's Rating. This gives users the ability to see CMS's Rating, while also, if they should so choose, incorporating their own values and preferences.

If CMS were to implement something similar to RAND's concept, the AAMC asks CMS to clarify how it would adjust the customized Ratings for a hospital where the hospital does not have a measure group score for a group a user has increased the weight to. In such a case, would CMS's tool recalibrate to give those weights to other groups included by the user? Or would it result in a "not applicable" response to alert the user that the hospital is unable to be measured on that user's criteria?

Another customization concept CMS could consider is one based upon condition. This would be a significantly bigger project to implement, as CMS would need to assess which conditions could be "singled" out with a rating with sufficient measures, and whether the condition-specific ratings are valid and reliable. Given the complexities of implementing this type of customization, the AAMC puts it forth only as a potential area to explore for the future as this may be the most meaningful type of information for many patients and consumers.

Generally, the AAMC is supportive of exploring user-customization to the Overall Hospital Quality Star Ratings. However, we believe more investigation is needed to better understand patient and consumer interest in and understanding of the concept to ensure that any tool meets their needs and is not overly burdensome or complex. Additionally, customized ratings must be reliable and valid before a customization tool is released to the public.

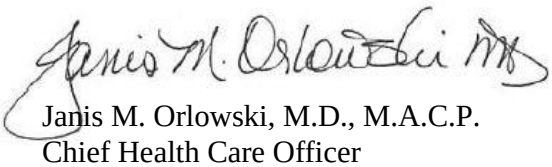
---

[10] See RAND Corporation "Personalized Hospital Performance Card" available at: https://www.rand.org/health-care/projects/personalized-hospital-performance-report-card.html (Visited March 25, 2019).

Conclusion

The AAMC welcomes engagement on these issues and appreciates the opportunity to comment. We look forward to continuing work with CMS on these issues. If you have any questions, please contact Gayle Lee at (202) 741-6429 or galee@aamc.org and Phoebe Ramsey (202) 448-6636 or pramsey@aamc.org.

Sincerely,



Janis M. Orlowski, M.D., M.A.C.P.
Chief Health Care Officer

Cc: Kate Goodrich, M.D., CMS
Reena Duseja, M.D., CMS
Ivy Baer, JD, MPH, AAMC
Matthew Baker, MS, AAMC