



American Hospital
Association®

800 10th Street, NW
Two CityCenter, Suite 400
Washington, DC 20001-4956
(202) 638-1100 Phone
www.aha.org

September 14, 2015

Patrick Conway, M.D.
Principal Deputy Administrator
Centers for Medicare & Medicaid Services
Department of Health and Human Services
Attention: CMS-1461-P
P.O. Box 801
Baltimore, MD 21244-8013

Re: *Comments on the Hospital Star Ratings Methodology Report*

Dear Dr. Conway:

On behalf of our nearly 5,000 hospitals, health systems and other health care organizations, the American Hospital Association (AHA) appreciates the opportunity to comment on the draft methodology for creating an overall star rating system for hospital quality as delineated in the report of CMS' Technical Expert Panel (TEP) and Aug. 13 webinar. Our members are dedicated to providing useful quality and safety information to our patients, the communities we serve and other stakeholders. **America's hospitals were instrumental in the creation of *Hospital Compare* more than a decade ago, and remain strongly committed to sharing meaningful, accurate hospital quality information with the patients they serve.**

It is critical that such information be presented in an understandable manner. As the Centers for Medicare & Medicaid Services (CMS) has noted, the data on *Hospital Compare* can be complex and difficult for many to understand and use. We support the concept of making this information more understandable, and recognize that the use of symbols, such as stars, may be an effective strategy to accomplish this goal. However, we have several concerns about the proposal under consideration.

GUIDING PRINCIPLES FOR DESIGN OF THE STAR RATING SYSTEM

CMS's guiding principles for creating a star rating system for hospitals are:

- **Simplicity and accessibility**, meaning CMS would summarize overall hospital quality in a single star rating and convey information in a straight-forward manner;
- **Inclusivity**, meaning that CMS would strive to include as many hospitals as possible in the star rating system;



- **Scientific rigor**, meaning CMS would use established methods for summarizing scores using existing data;
- **Engage stakeholders**, meaning CMS would seek input on the process throughout its deliberations; and
- **Consistency**, meaning CMS would align the hospital star rating system as much as possible with its other provider rating systems and use measures that are already part of other hospital measurement programs.

The AHA appreciates CMS's articulation of its guiding principles. We are particularly supportive of the agency's desire to include as many hospitals as possible in the star rating display, ensure the methodology is scientifically rigorous, and provide opportunities for stakeholder engagement.

However, consumers, other users and the hospitals that are being rated will have to rely on CMS to portray information fairly and accurately, make reasonable judgements about which measures are important and reliable enough for inclusion and which measures are relatively more important than others in aggregating everything into a single star rating. These ratings are intended to drive patients to wise choices about where they seek care. Thus, CMS has an obligation to ensure the differences in ratings it portrays are real, meaningful and important to patient outcomes. Users and hospitals have a reasonable expectation that, if CMS is going to assign star ratings to hospitals for the purpose of identifying different levels of performance, CMS will be able to substantiate its assertion that a three star hospital is more likely to deliver care patients would find superior to that delivered at a one- or two-star hospital, and less likely to deliver superior care than a four- or five-star hospital.

Ideally, this work would have begun by identifying a small set of scientifically sound measures that broadly look at critical aspects of care and that are measures consumers find compelling when choosing a hospital. But, that is not the basis upon which CMS chose the measures that appear on *Hospital Compare*. Instead, CMS has a set of measures, some of which were created to meet specific legislated program needs, like the readmission measures, and others that were created for a variety of research and registry purposes. As described in the report of the TEP discussions, CMS is trying to use this convenience sample to construct a meaningful star ratings approach. It is a daunting task.

Yet, the measures on *Hospital Compare* largely focus on important but relatively narrow aspects of hospital services like care for heart attack, stroke or pneumonia. There are only a small number of measures on *Hospital Compare* that reflect cross-cutting issues affecting many patients. The measures themselves range from patients' assessments of the cleanliness and quiet of the inpatient care facility, to measures of key outcomes that apply only to Medicare fee-for-service patients with particular conditions, to measures of how efficient hospitals were in providing certain imaging services in the outpatient setting. It is not clear to what extent, if any, these are the measures that would be most relevant to patients or other users if they were to describe what they would want to be incorporated into a star rating system of hospitals.

The AHA previously commented on the measures proposed for use in the hospital star rating system, suggesting that because individuals seek care at hospitals for a variety of different reasons, and that because hospitals' performance for patients with different conditions can vary based on many factors, it is unlikely that **a single summary star rating would equip patients, families and communities with a meaningful, accurate picture of hospital quality that is relevant to their individual reasons for seeking care. Moreover, we were not confident that the measures available on *Hospital Compare* would enable CMS to create a single, methodologically sound rating of all aspects of hospital quality. Thus, we urged the agency to consider a star rating system in which it applies star ratings only to specific measure topics, like cardiac care, rather than one overall rating for each hospital.**

Since then, an important article and editorial from the BMJ journal *Quality and Patient Safety* have come to our attention. The article by Paddock, et al titled "Better-than-average and worse-than-average hospitals may not differ significantly from average hospitals" explores the accuracy of using CMS readmission measures to differentiate performance among hospitals. The accompanying editorial by David Shahian and Sharon-Lise Normand, two researchers from Harvard, notes that, in fact, the readmissions measures CMS uses cannot be used to compare the performance of one hospital to another. Rather, the way the measures are constructed, they compare the performance of each hospital to a hypothetical "average" hospital with the same case mix as the original hospital. The rate that is calculated for each hospital is a comparison of actual to "expected" performance for that unique case mix, and is not comparable to the performance rate for a different hospital with its own unique patient population mix. Or as Shahian and Normand put it, "...while the CMS website is named *Hospital Compare*, the statistically valid comparison is between each hospital and a hypothetical average hospital, not between pairs of hospitals." The same methodological concerns exist for the mortality measures, which are constructed using the same methodology of comparing a hospital to a hypothetical average hospital with the same case mix and average performance for that case mix.

Thus, we believe it is not possible to accurately assign star ratings to hospitals based on these non-comparable performance measures. The measures are simply not constructed in a way to permit this kind of hospital-to-hospital comparison. **We urge CMS to reconsider using readmission and mortality measures in the star rating system since they will likely lead to the misclassification of hospitals, resulting in misinformation for patients.**

LATENT VARIABLE MODEL

CMS proposes to group existing measures into seven categories – mortality outcomes, safety outcomes, readmissions outcomes, patient experience, timeliness of care, effectiveness of care and imaging efficiency. CMS will then apply a latent variable model to each of these groups. This model assumes that there is a single common factor that is present for the measured performance of a hospital on each of the measures in a group, and that factor could be described as the hospital's influence on the measured performance. CMS proposes to use this analysis to determine the value of the "latent variable" and then use the latent variable as the measure of how a hospital performed on each category. A weighted average of all of the hospital's latent variable factors will then be used to determine how many stars a hospital receives.

While statisticians might appreciate this approach, using the latent variable model would mean that hospitals will not recognize the measured performance in any category as consistent with the data on *Hospital Compare*. Hospitals can legitimately say that the number is not consistent with how they scored on any of the measures, nor is it an average of those scores. Further, they will have no idea what to do to improve their performance on this latent variable. It is an interesting statistical method for dealing with different measures and highly varied performance on measures, but it will certainly come across as an inexplicable, “black box” assessment to many users, including many hospital leaders. More importantly, it may stifle improvement efforts since there will not be a clear path toward improving the score. **We urge CMS to take a step back and consider something simpler and with a more direct “line of sight” between measured performance and performance improvement actions.**

WEIGHTING OF MEASURE CATEGORIES

To create an overarching star rating from among the performance scores calculated for the seven different measurement groups, the report proposes weighting four of the seven categories at 22 percent and three at 4 percent. As in the value-based purchasing program, CMS proposes heavier weights for the outcome measure categories and the patient experience category, and smaller weights for the process measures and the imaging efficiency category. There is no real justification offered for the weights chosen for each of the categories.

These weights are actually critical to the determination of how many stars are received. **We believe that it is imperative that CMS justify the weights for each category based on the scientific integrity of the measures and the preferences of patients about how important that category is to their assessment of the hospital performance. We urge CMS to give this weighting much more attention to ensure it is justifiable.**

Further, the report notes that not all hospitals will be reporting all measures used in each of the categories, and some may have entire categories of measures missing. A number of options appear to have been explored with regard to determining the minimum number of measures a hospital must have to get a score in any group and the minimum number of groups in which a hospital must have a score in order to get a star rating. The recommendation seems to be that every hospital getting a star rating have scores in at least three groups, one of which must be an outcome group, and that, to get a score in a group, there must be at least three scored measures.

We are concerned about the likelihood that these small numbers of required categories and measures will result in misleading assignments of stars. These proposed minimums would mean that CMS could be assigning stars to hospitals based on categories to which only 30 percent of the weight would otherwise have been assigned. We believe the question of how many measures need to be in each category should be answered empirically rather than by guesswork or voting. The answer should be determined by how many measures are needed to provide a relatively stable estimate of what the hospital’s score would have been if all factors had been present, and the answer may vary by category. It seems much more likely that three measures would provide

a representative sample out of the six in the mortality measure category than three of the 11 in the patient experience category or three of the 30 in the effectiveness of care category.

In conclusion, we support both the concept of providing an easier way for patients and community members to understand what the published measures are telling them about the quality of care being provided in America's hospitals. We also support the majority of CMS's principles for creating a star rating system and are glad to assist in thinking through how to overcome many of the impediments to successfully creating such a system. However, we are concerned that CMS is trying to retrofit the existing measures into telling patients and consumers something meaningful when those measures were not chosen with this goal in mind, and many are not constructed in such a way that they can accurately depict the kind of distinctions in care CMS wants to depict.

We urge CMS to consider moving forward with the goal of a creating a comprehensive assessment of hospital care within the next few years rather than all at once. The Agency can build from its initial star ratings based on HCAHPS, and add condition specific assessments. This will provide an opportunity to strategically add measures that can provide important insights into overall performance as well as take advantage of conditions for which we currently have enough information to provide an accurate assessment. We can envision component pieces, such as a star rating safety, or on some of the specific diagnostic conditions for which CMS currently has a sufficient number of measures serving as the early building blocks toward the more comprehensive star rating system.

The Institute of Medicine articulated a framework for getting to more meaningful measurement in its *Vital Signs* report. We believe that, with leadership from CMS and the Department of Health and Human Services, it is highly likely that this framework can help to get us to a point where we have measures that will support the meaningful star rating system CMS intends to create.

Please feel free to contact Nancy Foster, vice president for quality and patient safety policy, at (202) 626-2337 or nfoster@aha.org if you have questions.

Sincerely,

Ashley Thompson
Acting Senior Executive for Policy

CC: Kate Goodrich, M.D.