

August 5, 2014

Dr. Patrick Conway
Chief Medical Officer
Director, Center for Clinical Standards and Quality
Centers for Medicare and Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244

Re: Addition of Star Ratings to Dialysis Facility Compare

Dear Dr. Conway:

We applaud CMS' well-intentioned efforts to try to simplify Dialysis Facility Compare (DFC) for consumers. However, we are not convinced that the proposed five-star methodology will accomplish our shared goal of improving transparency for beneficiaries. It seems to us that the agency is moving very quickly to adopt a novel scoring methodology that exposes patients to two conflicting quality rating systems for facilities, and is doing so at a time when consensus has formed in the health policy community that outcome measures need to be adjusted for socio-economic status (SES). We would like some reassurance that when CMS made the decision to apply the star ratings to Dialysis Facility Compare, the question considered by agency officials was whether the proposed system would enhance the knowledge and engagement of dialysis patients, and not simply how to implement a policy of expanding star ratings to all categories of providers.

We urge CMS to take any additional time necessary to consider (1) whether the proposed bell-curve scoring mechanism presents information in a way that is helpful to consumers; (2) whether the star ratings are appropriately aligned with the ongoing Quality Incentive Program; and (3) whether the initiative should proceed in a nationwide tournament format without further accounting for differences in the underlying health of populations served by particular facilities. We would appreciate an opportunity to work with the agency on refinement of this project and offer to assist in recruiting patients for focus groups to comment on the presentation formats prior to deployment.

Because we anticipated the introduction of a five-star rating for DFC, we added a question to our annual Member Survey asking patients whether they felt giving star ratings to dialysis facilities would improve their understanding of quality of care. Eighty-five percent of respondents answered in the affirmative. Further, 72 percent of respondents said they would be very likely

(44%) or somewhat likely (28%) to consider switching facilities. These responses support creation of a star program, but we believe they also underscore the necessity of launching the program with caution and regard for patients' reasonable expectations.

While there is a broad consensus favoring transparency in health care, a number of scholars have documented difficulties in the mechanics of presenting quality measures and disclosures to patients. As noted in a recent white paper on communication of quality measures produced by a Harvard School of Public Health team, "while making quality information publicly available online is relatively simple, presenting the information in a way that is useful to consumers can be a significant challenge."¹ Below we describe what we view as barriers to consumer understanding of the five-star program as it has been described thus far.

1. Concerns Relating to Bell Curve Scoring Methodology

If there is one theme that runs throughout the scholarly literature on consumerism in health care, it is the importance of giving patients proper *context* for the information that is being presented. In this instance, our concerns about the context of the bell-curve scoring arise from the unique nature of dialysis and from consumers' prior experience with star rating scales. The sphere in which consumers most frequently see star scales is in reviews of discretionary purchases such as movies, restaurants or lodging. In these circumstances, a judgment of two or fewer stars is generally understood as advice to *not* make the purchase, e.g. don't see this movie, don't eat at this restaurant. Until now, CMS has used star rankings for discretionary health care purchases, such as Part C health plans (to which fee-for-service Medicare is an alternative) or nursing homes (to which remaining in the community may be an alternative). If the beneficiary's only options are health plans or nursing homes with one star, he or she can pursue other avenues of receiving care.

Dialysis is not a discretionary purchase—it is necessary for a person with kidney failure to stay alive. So we wonder what reaction a patient will have to finding that his or her only nearby options for dialysis are facilities with one or two stars. We suspect that this may be the case in certain regions where historically poor performance on outcome measures will likely result in facilities falling in the bottom 30 percent. We do not have access to the star measures, but we can infer from publicly available information that one such area might be rural Louisiana. Opelousas, Louisiana is a majority African-American town where 43% of the population lives below the poverty line. For a dialysis patient residing in Opelousas, there are eight facilities within 26 miles. Of these eight, three have worse than expected mortality, and only one, located 24 miles away, has a standardized mortality ratio (SMR) below one. The other seven have SMRs of 1.28 or greater. It is not clear to us precisely what our hypothetical Opelousas patient is supposed to do upon being informed that all of his or her nearby options are one- or two-star facilities.

¹ Viswanath K et al. Communication and Quality of Care: An Overview.
http://jktgfoundation.org/data/QOCInformationWhitePaper_2014July%20FNL.pdf

As Tracey Miller and William Sage have noted, disclosures about health care providers must strike the proper balance “between educating patients and alarming them.”² Given the connotations commonly associated with one- and two-star ratings in the minds of consumers, we are concerned that these ratings may inappropriately stigmatize facilities for outcomes that are beyond their control. We hope the patient would not interpret these reviews as one ordinarily interprets a movie review, and decide to stay home. We further note that for beneficiaries who regularly use the consumer review website Yelp, one- and two-star facilities may be judged particularly harshly—Yelp issues those ratings to just the lowest *twenty* percent of businesses.

We certainly agree with the premise of bell curve scoring—that our most cherished beliefs to the contrary, *everyone* is *not* above average. But it is not clear how the 30 percent threshold for negative scores was selected, nor why a relatively high cutpoint would be put in place for facilities whose use, unlike a movie exhibition or restaurant meal, will rarely be a waste of the consumer’s time or money. It seems to us that the safest course would be to limit one- and two-star ratings to those facilities for which CMS has a high degree of certainty that poor outcomes are the result of substandard clinical practices or management, as revealed, for instance, by an inspection.

Finally, we would like to see CMS’ analysis of the number of patients currently in one- and two star facilities who would have reasonable access to a three-star or better facility (e.g., a three-star facility within 10 miles or 20 minutes), as well as the number of patients in “average” facilities who could upgrade to an “above average” facility. If a significant majority of patients could not realistically act upon the star ratings, we are not certain that this information would be helpful.

2. Concerns Relating to the Alignment of Star Ratings with QIP

The patient’s processing of the star ratings is further complicated by conflicting information in the performance score certificates (PSC) for these facilities. We know of at least one facility in Eastern Kentucky that, while having worse than expected mortality and hospitalization rates, has higher than average scores on all four of the clinical measures of quality in the Quality Incentive Program, which means that the certificate in that facility is marked with “Yes” in all four boxes under “Meets Standard.”

As you know, many providers have complained about receiving different quality ratings from different payers, and there is general bemusement about variations in rankings bestowed by non-payer entities, giving rise to one-liners like “fifty of the twenty-five best hospitals are in New York.” However, we believe this is the first time that patients have faced the same payer issuing more than one rating to the same health care provider. (Note: According to our patient survey, 11 percent of patients have visited the Dialysis Facility Compare website, and 41 percent have seen the QIP poster in their facility.)

² Miller TE, Sage WM. Disclosing Physician Financial Incentives. *JAMA*.1999;281(15):1424-1430. doi:10.1001/jama.281.15.1424.

We are confident that CMS staff has anticipated circumstances such as this, so we are not going to pre-judge whether two rating systems standing side-by-side are inherently incompatible. But we are curious as to how the agency sees the very different disclosures presented on DFC and the PSC fitting into what Miller and Sage refer to as an “integrated communication strategy.” We hope you will share with us the rationale, and would also appreciate an opportunity to preview any draft language that has been proposed to guide beneficiaries in weighing the relative importance of the two disclosures when they diverge.

3. Concerns About Nationwide Comparisons

From the beginning of its quality reporting and pay-for-performance efforts, CMS has taken the position that quality should be judged as part of a nationwide competition among providers. This approach assumes that providers serving the most disadvantaged areas in our very diverse nation can, or at least should, be capable of producing as favorable patient outcomes as their counterparts in wealthier regions. That assumption has been challenged recently by the Medicare Payment Advisory Commission (MedPAC), the National Quality Forum (NQF), and in bipartisan, bicameral legislation currently pending in Congress. MedPAC and an NQF panel have called for adjusting measures for socio-economic status, and the pending legislation would require CMS to do so.

The criticism of CMS’ position has largely been driven by experience with the hospital readmissions penalty. The most severe penalties appear to be assessed against urban safety-net hospitals and hospitals in regions with poor population health. For instance, of 18 hospitals receiving the full 2 percent penalty, ten are in Greater Appalachia, three are in Texas, and two are in Louisiana, with only two north of the Mason-Dixon line. The pattern of readmissions, despite risk adjustments, closely resembles the maps produced by Christopher Murray and Majid Ezzati that depict mortality by county and race to carve out what those researchers dubbed “Eight Americas”—distinct American subpopulations that are either favored or disfavored in terms of health outcomes.³ Our government affairs director, Jackson Williams, presented his research to you two years ago regarding similarities in the Ezzati/Murray maps to a map of U.S. Regional Subcultures produced by Joel Lieske. The Williams paper identifies a pattern in which cultural characteristics of a subpopulation, beyond SES, seem to drive health behaviors such as medication adherence, leading to differential outcomes.⁴

We note that the distribution of dialysis facilities with “worse than expected” mortality closely corresponds to these geographic patterns. In Kentucky, 12 facilities have “worse than expected” mortality while only one has “better than expected” mortality. In West Virginia, 7 facilities have “worse than expected” mortality while only one has “better than expected” mortality. These states lie in what Murray calls “America 4,” an area where life expectancy is “similar to those of

³ Murray CJL, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, et al. (2006) Eight Americas: Investigating Mortality Disparities across Races, Counties, and Race-Counties in the United States. *PLoS Med* 3(9): e260. doi:10.1371/journal.pmed.0030260

⁴ Williams J. Regional cultures and health outcomes: Implications for performance measurement, public health and policy. *Social Science Journal* 01/2013; 50(4):461–470.

Mexico and Panama” and which bore the most 2% readmission penalties. In Louisiana there are 24 facilities with “worse than expected” mortality and 4 “better than expected” mortality. Alabama has 14 facilities with “worse than expected” mortality and 4 with “better than expected” mortality. These states lie in what Murray identifies as “America 7,” comprised of “low-income rural blacks in the Mississippi Valley and the Deep South,” which generally has the highest mortality rates of any of the eight demographic classifications.

The reverse of this pattern is evident in regions known for good health. Together, Minnesota and Wisconsin, dominated by Murray’s “America 2,” have 20 dialysis facilities with better than expected mortality and just 7 with worse. Note that Murray et al describe America 2 as dominated by “low-income rural white populations, with income and education below the national average,” emphasizing that SES is a necessary but insufficient adjustment. Turning to regions that Joel Lieske classifies as “Rurban,” we see that in Colorado, 13 dialysis facilities have better than expected mortality and only one worse; in Washington State, 6 facilities are better and one is worse. CMS personnel who have pointed to Denver Health as having good performance on readmissions despite being a safety net hospital should carefully examine the Murray maps of life expectancy by race, which show that African Americans in Denver have high life expectancy relative to those in other parts of the U.S. Murray et al do not include Denver in their “America 8” which consists of “blacks living in high-risk urban environments.”

The doubt that these factors cast on the validity of quality measures is illustrated by circumstances in Opelousas. As noted, seven of the nearest dialysis facilities have high mortality, despite being operated by different personnel, reporting to different managers, with different medical directors and subject to control by two different large dialysis organizations with different policies and procedures. Meanwhile, the four IPPS hospitals nearest to Opelousas have received readmissions penalties, two of them rather stiff. These hospitals of course have different personnel and management than the dialysis facilities, yet their patients, too, have poor outcomes. It strikes us as a highly unlikely coincidence that eleven different providers with eleven different staffs but serving similar populations would all show poor outcomes solely because their clinical skills are similarly substandard. What seems more likely is that quality measures are conveying information about the poor population health in the region: St. Landry Parish has a premature death rate (years of potential life lost before age 75 per 100,000 population) that is double the national average, and nearly triple the rate of Boulder County Colorado, against whose dialysis clinics those in Opelousas must compete for stars.

We believe that MedPAC has pointed to the correct solution to this problem: instead of holding a national competition, CMS could cluster facilities serving similar patient populations into “peer groups” for quality comparison purposes. In such a regime, the facilities in Louisiana could be judged against each other, not against counterparts in Colorado or Minnesota that set seemingly unattainable standards for them. Clinicians would have to step up their game in every region, because competition would be realistic and the strength of opponents would be no excuse for falling behind. Peer grouping for measures means no reputational punishment for serving disadvantaged communities, so there would be no incentive for national large dialysis organizations to divest facilities in low-income regions to maintain a higher average star rating for their chains as a whole.

We request that CMS release a map showing the geographic distribution of the star ratings under the bell curve methodology for us to review prior to the go-live date.

4. We Urge a Collaborative Approach to Moving Forward

We were pleased to hear that CMS conducted consumer testing of the five-star format. But given that consumer testing assesses the accuracy of patients' interpretation of measures, the agency must explicitly or implicitly make a normative judgment as to what interpretation is "accurate" and what the consumer is intended to understand after being exposed to the presentation. It would be helpful for us to know:

- Are the one- and two-star ratings intended to be cautionary—i.e., are consumers being told to avoid low-scoring facilities if possible?
- What are patients being advised to do if only low-rated facilities are available in the vicinity of their homes?
- When the star rating conflicts with the Performance Score Certificate, which impression of the facility was deemed "accurate" in testing scenarios?
- When Nursing Home Compare was being tested by Mathematica, some respondents had questions about whether ratings were affected by sicker patient populations. Did this occur in testing DFC, or did CMS anticipate this beforehand? How has CMS dealt with this issue?

It is easy for patient advocates to be skeptical about new formats for presenting quality information to consumers that have not been subject to a notice-and-comment dialogue with stakeholders. As Mathematica noted in its debrief on the Nursing Home Compare testing:

Quality information may be technically complex. Consumers accustomed to thinking of health care in personal terms may not understand how aggregate measures of performance... relate to them. They may not be aware of systematic variations in quality, or they may perceive that they have little choice, in any case. Disturbingly, the research reported here suggests that even when they are engaged, consumers may erroneously interpret quality information without knowing that they are doing so.⁵

We do not see the need for any rush to roll this out on an arbitrary timetable, nor a need to change DFC identically to or staggered with other "Compare" websites to which star ratings are being added. We also see this as an opportunity to begin thinking about how to devolve inter-provider quality competition to appropriate subnational peer groupings, or at the very least, an opportunity to avoid further expansion of nationwide tournaments at a time when a broad cross-section of health policy thought leaders is questioning that approach.

⁵ Gerteis M, Gerteis JS, Newman D, Koepke C. Testing consumers' comprehension of quality measures using alternative reporting formats. *Health Care Financ Rev.* 2007 Spring;28(3):31-45.

Unlike hospitals and home health agencies, dialysis facilities have a base of identifiable, longitudinal patients who are organized under the auspices of organizations like ours and who are available to volunteer for focus groups or other methods of providing feedback on proposed quality measure presentation formats. We hope you will give us the opportunity to work with the Agency on this project.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Hrant Jamgochian".

Hrant Jamgochian, J.D., LL.M.
Executive Director

A handwritten signature in black ink, appearing to read "Jackson Williams".

Jackson Williams
Director of Government Affairs