Inpatient Quality Reporting (IQR) Program

Support Contractor

Overall Hospital Quality Star Ratings on *Hospital Compare*

Questions & Answers

Moderator:

Candace Jackson, RN
Project Lead, Hospital IQR Program
Hospital Inpatient Value, Incentives, and Quality Reporting (VIQR)
Outreach and Education Support Contractor (SC)

Speaker:

Pierre Yong, MD, MPH, MS
Acting Director of Quality Measurement and Value-Based Incentives Group (QMVIG) in the Center for Clinical Standards and Quality at the Centers for Medicare & Medicaid Services (CMS)

Arjun Venkatesh, MD, MBA, MHS
Assistant Professor and Director of Quality and Safety Research and Strategy in the Department of Emergency Medicine at the Yale University School of Medicine

Kristie Baus, RN
Technical Advisor, CMS
Technical Lead, *Hospital Compare* Website

May 12, 2016
1 p.m. ET

APPEAL PROCESS

Question 1: **How can hospitals correct incorrect information on the *Hospital Compare* website? For example: an error in the "provides emergency services" tab (e.g., a facility that provides emergency services displays as "no," while one that does not is accidentally listed as "yes")?**

Please submit a request, including the hospital's CMS Certification Number (CCN) via the Inpatient Questions and Answers tool at: https://cms-ip.custhelp.com or the Outpatient Questions and Answers tool at https://cms-ocsq.custhelp.com

Question 2: **What is the Appeals Process for a Hospital to challenge its rating? It may think it deserves a 3-star rating and CMS gives it a 2-star rating.**

Please submit a request, including the hospital's CMS Certification Number (CCN) via the Inpatient Questions and Answers tool at: https://cms-ip.custhelp.com or the Outpatient Questions and Answers tool at https://cms-ocsq.custhelp.com

BEHAVIORAL HEALTH/INPATIENT PSYCHIATRIC FACILITIES

Question 3: **Is it mandatory that Behavioral Health/Psych hospitals participate? How are they currently rated by CMS?**

No, the Star Ratings are only applicable to hospitals participating in the Hospital IQR and Outpatient Quality Reporting (OQR) Programs.

CRITICAL ACCESS HOSPITALS (CAHs)

Question 4: **What is the requirement for CAHs to report Web-Based measures? Are they included in the star rating?**

Answers to this question can be found on the QualityNet *website. Please visit the website to learn how hospitals can participate.*

Question 5: **Web-Based measures are used in the Methodology as long as they are excluded by the measure selection criteria?**

CMS uses the following criteria to exclude measures from the Star Ratings calculation:

1. *Measures suspended, retired, or delayed from public reporting on* Hospital Compare*;*
2. *Measures with no more than 100 hospitals reporting performance publicly;*
3. *Structural measures;*
4. *Measures for which it is unclear whether a higher or lower score is better (non-directional);*
5. *Measures no longer required for the IQR or OQR Programs; and*
6. *Duplicative measures (e.g., individual measures that make up a composite measure that is also reported; or measures that are*

identical to another measure).

CATEGORICAL GROUP PERFORMANCE

Question 6: **Is the categorical group performance going to be presented in the hospital reports or the publicly reported data on *Hospital Compare*?**

CMS is working with stakeholders to determine if this information is useful for consumers. If we do decide to make this information public, it will be done in the downloadable database rather than the Hospital Compare *workflow.*

Question 7: **Have hospitals received their preview information yet with the categorical group scores that were referenced on slide 49?**

The July 2016 Methodology report, Frequently Asked Questions (FAQs) and Quarterly Updated Specification reports will be posted on Quality Net *soon*

CHILDREN'S HOSPITALS STAR RATINGS

Question 8: **We've noticed that many children's hospitals do not have complete data profiles on *Hospital Compare*. Are there any plans for new measures so that children's hospitals can have star ratings similar to other hospitals?**

The law states that only subsection d acute care hospitals paid under the Inpatient Prospective Payment System (IPPS) or Outpatient Prospective Payment System (OPPS) are at risk for a payment reduction in their annual payment update if they do not participate in the IQR or OQR Programs. Many facilities that do not fall under those categories, including children's hospitals and CAHs, voluntarily submit data applicable to their facilities for public reporting. They are not required to submit data, and are not subject to any payment reductions for not submitting data.

Support Contractor

CLAIMS DATA

Question 9: **Why is the claims data publicly reported unavailable to the facility to review?**

The Overall Hospital Quality Star Rating represent a summary of performance based on specific measures currently available on Hospital Compare *and does not provide patient-level data. However, the Hospital IQR and OQR Hospital-Specific Reports (HSRs) provide patient-level data.*

Please contact the Hospital IQR support contractor at iqr@hsag.com or 844.472.447/866.800.8765, or the Hospital OQR support contractor at oqrsupport@hsag.com or 866.800.8756, for any additional questions regarding the program-specific HSRs.

COMMENTS

Listed below are a series of comments provided by some of our webinar participants. CMS thanks you all for your comments.

Question 10: **Since the mortality metric has been standardized and now has a reverse direction, should the name be changed to "Survival" rather than "Mortality?" My institution as 'above average' mortality, which on the surface, sounds bad. But if we had above average survival, that would sound good.**

Thank you for that suggestion. It is something that the workgroup has discussed and we will definitely take that under consideration.

Question 11: **There will continue to be controversy about the outcome measures until sufficient and robust risk adjustment for sociodemographic factors is included in the methodology for calculation. (For example the methodology of Kind et al. Ann Intern Med. 2014 Dec 2; 161(11):765-74.)**

Question 12: **CMS has already instituted a "star" rating for HCAHPS; the inclusion of this in the overall "star" rating is double jeopardy and unfairly skews either positively or negatively.**

Support Contractor

Question 13: **Neither TJC nor Health Grades ratings models role up to a single score. We understand your statistical methodology, but it is irresponsible on the part of CMS to imply that this roll-up number accurately reflects the overall performance.**

Question 14: **The weight redistribution to non-missing measure groups has the potential to put hospitals with all groups at a disadvantage. For example, small hospitals consistently perform better on Patient Experience Measures and are less likely to have sufficient volume to be evaluated for the Outcome Measures. So, for those hospitals, Patient Experience contributes more to their overall rating.**

Question 15: **In reference to public perception, would CMS consider putting a general disclaimer underneath the star rating, such as "This star rating may not reflect current performance due to inclusion of data from older time periods."**

Question 16: **If you want to use PSI-90, it should be noted on the public website that these are derived from an Administrative Data set.**

Question 17: **Several of the measures only include Medicare patients, so to label it as an overall Star Rating is misleading. Key groups are not included (i.e., Perinatal Care and Pediatrics).**

Question 18: **CMS' "Easily understood star rating" is IMPOSSIBLE to reproduce.**

Question 19: **NHSN measures include hospital characteristics which are used in the CMS stars, while other measures do not. This seems like an inconsistent approach. I would request CMS provide a consistent message approach.**

Question 20: **Please provide hospitals with a tool to assess the impact of a given measure on our star rating.**

eCQMs

Question 21: Will the performance on the four eCQMs required for 2016 IQR be included in future star ratings?

The star ratings are designed to be as inclusive of as many measures posted on the Hospital Compare website. So, at a time when CMS does decide to add the results of the eCQMs, we'll determine then whether or not they'll be included in the star ratings.

Question 22: When eCQM are implemented how soon will they be incorporated into star ratings?

CMS will evaluate new measures that are added to Hospital Compare *using the measure inclusion criteria developed with stakeholders and experts through Technical Expert Panel (TEP) meetings and public comment.*

Question 23: Will eCQM measures be added in the future?

CMS will evaluate new measures that are added to Hospital Compare *using the measure inclusion criteria developed with stakeholders and experts through TEP meetings and public comment.*

Question 24: How will the transition to eCQMs affect the Effectiveness of Care Measures that have been moved to electronic abstraction? Many of the rates are not reflective of actual care provided.

Once the eCQMs are implemented for public reporting, CMS will have a discussion regarding incorporating the eCQMs into the Overall Hospital Quality Star Rating in a future release.

Question 25: Will an actual list of removed measures be available?

Please refer to the July 2016 Quarterly Update Specifications Report found on Page 18, Table B.2. Measures Excluded from July 2016 Star Ratings located on QualityNet.

Question 26: **Can we get a list of the measures that were excluded? Slide 17**

Please refer to the July 2016 Quarterly Update Specifications Report found on Page 18, Table B.2. Measures Excluded from July 2016 Star Ratings located on QualityNet.

HCAHPS

Question 27: **Any thoughts on why all 5-star hospitals in the HCAHPS summary are under 400 beds, predominantly for-profit, and suburban? Can we expect a similar distribution in the overall star ratings?**

We are currently doing some analysis on the distribution of the star ratings, and those results will be forthcoming at a later day.

The HCAHPS Project Team is available for discussion of HCAHPS performance. There were no surprises in the HCAHPS star ratings compared with the HCAHPS measure scores.

Question 28: **Will HCAHPS stars be phased out, or in 2017 will they be updated quarterly while star ratings are refreshed semi-annually? How will we know what time periods are used for the measures for each refresh?**

HCAHPS Star Ratings will not be phased out. HCAHPS Star Ratings will be updated quarterly on Hospital Compare*, which includes information about the time period covered by the HCAHPS Star Ratings and other HCAHPS measures.*

Question 29: **Are the HCAHPS Star Ratings (which are also a part of Overall Star Ratings) available publically for May 2016 *Hospital Compare* Refresh?**

Yes, the HCAHPS Star Ratings on Hospital Compare *were refreshed in early May.*

Question 30: **For the weighting of the HCAHPS group score for each hospital, I can calculate the z-score across the 11 metrics, but I need to know the weighting of each of the 11 to calculate the group score, correct?**

The weighting of the 11 HCAHPS measures in the HCAHPS Summary Star

Rating is explained in the HCAHPS Star Rating Technical Notes available at http://www.hcahpsonline.org/StarRatings.aspx.

HOSPITAL COMPARE UPDATES

Question 31: Is there a document showing which *Hospital Compare* metric is scheduled to be updated with each release?

Please submit your inquiry to the Hospital Compare *technical assistance inbox at HospitalCompare@hsag.com for assistance.*

HOSPITAL-WIDE READMISSION

Question 32: Why is Hospital-wide Readmission included if CMS specifically excludes? Duplicative measures?

The measure is endorsed by The National Quality Forum (NQF) and has undergone rigorous testing for scientific acceptability and validity. CMS believes that the measure is an important indicator of overall hospital quality. CMS developed the Overall Star Rating to be as inclusive of as many measures currently reported on Hospital Compare *as possible in order to present the most comprehensive picture of hospital quality for consumers. The Overall Star Rating methodology includes a systematic process for determining the eligibility of a measure for inclusion which was vetted by the TEP and public comment. In addition, by studying the star ratings data, CMS concluded that it is unlikely that any one measure precludes a given type of hospital from performing well. For example, a hospital that has poor performance on a single Safety of Care measure, such as PSI-90, may still receive a high Safety of Care group score and a high star rating if that hospital performs well on the other included Safety of Care measures. Similarly, CMS does not believe that the removal of the Hospital-Wide Readmission measure would materially change a hospital's Overall Star Ratings results.*

HOSPITAL-SPECIFIC REPORTS (HSRs)

Question 33: Will there be a hospital-specific report showing all the detail scoring?

Yes, hospitals will receive HSRs for the July 2016 release. The HSRs will include Preview Report data, as well as Confidence Intervals for the group

score and summary score, and standardized scores for the measures. HSRs for IQR Claims-Based Measures are also being released. Both sets of reports are now available.

Question 34: How do we access the preview reports and HSRs and determine who in our organization they are being sent to.

The Hospital Compare *preview reports are available from May 6 through June 4, 2016. Hospitals are encouraged to access and download reports early in the preview period in order to have time for a thorough review. The preview reports and the ability to download certain reports are only available during the preview period.*

Facilities may access their preview reports by logging in to the QualityNet Secure Portal *and selecting the report they wish to view. The preview report data will be reported on CMS'* Hospital Compare *website where Medicare beneficiaries and the general public can review data indicators on quality of care for participating hospitals and facilities.*

Reports can be previewed by:

- *Accessing the public website for* QualityNet *at* https://www.qualitynet.org
- *Selecting* ***[Login]*** *under the Log in to* QualityNet Secure Portal *header*
- *Entering your* QualityNet *User ID, Password, and Security Code and selecting* ***[Submit]***
- *Reading the Terms and Conditions statement and selecting* ***[I Accept]*** *to proceed*

Preview Reports can be run by selecting:

- ***[Run Reports]*** *from the My Reports drop-down*
- ***[IQR]*** *or* ***[OQR]*** *from the Report Program drop-down*
- ***[Public Reporting – Preview Reports]*** *from the list in the Report Category drop-down*
- ***[View Reports]*** *– the selected report will display under Report Name*
- ***[Public Reporting – Preview Reports]*** *under Report Name*
- ***[Run Reports]***

Selected Preview Reports can be viewed and downloaded by:

- *Selecting the* ***[Search Reports]*** *tab*
- *A green check mark will display in the Status column when a report is complete*

Question 35: **How will the HSRs be released and how will we be notified of their availability?**

Yes, hospitals will receive HSRs for the July 2016 release. The HSRs will include Preview Report data, as well as confidence intervals for the group score and summary score, and standardized scores for the measures. HSRs for IQR Claims-Based Measures are also being released. Both sets of reports are now available.

Question 36: **It is mentioned that the release date of HSRs on Star Rating is to be determined. Do you have a tentative time line, such as what months this year?**

Yes, hospitals will receive HSRs for the July 2016 release. The HSRs will include Preview Report data, as well as confidence intervals for the group score and summary score, and standardized scores for the measures. HSRs for IQR Claims-Based Measures are also being released. Both sets of reports are now available.

IMAGING EFFICIENCY MEASURES

Question 37: **Please verify. Did you say Imaging Efficiency Measures will not be used to determine star ratings in the July report?**

The Efficient use of imaging measures are included in the star ratings calculation for July 2016. For a list of included measures and the measure groups, please see the methodology resources located on QualityNet.

Question 38: **The imaging efficiency measures tend not to have a set rate that indicates optimal care. We are told that generally, a lower number is better. However, we were cautioned that too low of a rate may mean that appropriate imaging is not being done. With this in mind, how can we group individual hospital rates and say that one group indicates better care than another group of hospitals?**

The measures that do not have clear direction have been removed from the star ratings methodology. Please see the included measures in the July Quarterly Update Report located on QualityNet. *CMS appreciates the feedback received regarding the measures in current publicly reporting.*

CMS will continue to evaluate all measures on Hospital Compare *as a part of current maintenance.*

Question 39: Why were the Imaging Efficiency measures included in the star rating system? They do not appear to be directly related to hospital efficiency.

Star ratings have a goal to reflect the current quality measures on Hospital Compare. *Existing measures may not capture "all" of hospital quality, however the current public reporting requirements result in diversity in the number and types of measures included in star ratings. The star ratings methodology holds the principles to be as inclusive as possible. The methodology also has measure exclusion criteria that has been vetted through a TEP and public comment. The efficient use of medical imaging measures does not meet the criteria to exclude from star ratings.*

INCLUDED MEASURES

Question 40: Do you have the detailed list of the 64 measures that are included?

Please refer to the July 2016 Quarterly Update Specifications report, Page 16, Table B.1. Measures Included in July 2016 Star Ratings (N=64).

Question 41: Is the data all Medicare FFS, or does it utilize all payer, Medicare Advantage, dual eligible or other categories?

The Claims-Based Measures, which include the Mortality, Readmission, Complications, PSI-90, and Imaging Efficiency measures are calculated using Medicare Fee-for-Service hospital claims data only. The Process of Care, HAI, and HCAHPS data is Chart-Abstracted and include information from all payers.

Question 42: When will the loadings and methodology reports be updated on *QualityNet* to reflect the time frames included in the July *Hospital Compare* data refresh?

The July 2016 Methodology report and Quarterly Updated Specification

reports are posted on QualityNet.

JULY PREVIEW REPORT TIMEFRAME

Question 43: What is the timeframe for the July report that will be reflected in the stars?

The measure dates for the Process of Care Measures data will be one quarter more – or advance one quarter – from where they are now. For example, it would be third quarter 2014 to second quarter 2015 data. The Process Measures, HCAHPS and HAIs, always roll forward a quarter.

For the majority of the Outcome of Care Measures in July, these are the measures that use three years of data. They will encompass July 1, 2013 to June 30, 2015.

LOADING

Question 44: Can you give a concrete example of the loading effect? For instance, in Safety, the PSI-90 has a much higher load than surveillance HAIs. How does this affect the group score?

A loading represents the association between an individual hospital measure and the group score. A measure with a higher loading value can be interpreted as having a higher correlation between that measure and the group score. In concept, your description of relative weight is correct. However, a loading is not a weight, or scalar proportion, in the conventional sense. Therefore, the relative weights you have calculated may be misleading because they cannot be used to calculate a simple weighted average of measure scores to generate a group score. Rather, the degree a given measure contributes to your hospital's group score is dependent upon the following:

- *Your hospital's measure score;*
- *Your hospital's measure denominator (case count);*
- *National performance on the measure; and*
- *The value of the loading relative to the loadings of other measures in the group.*

If you perform well on a measure with a large denominator (indicating greater precision of the measure score estimate), broad distribution of national performance, and high loading, this measure will contribute more to

your group score than a measure for which any of these characteristics are reversed. In other words, if a measure has the same loading value but a narrow distribution of national performance, this measure will contribute less to your group score. Therefore, the loadings alone cannot be used to evaluate the measure's relationship to the group score. We encourage you to evaluate the loadings within the context of your hospital's measure score and national performance on each of the measures in the group.

The statistical equation for the latent variable model can be found on page 14 of the 2015 Dry Run methodology report located on QualityNet *at www.qualitynet.org > Hospitals – Inpatient > Hospital Star Ratings > Previous Resources > "Overall Hospital Quality Star Rating Methodology Report" (fifth bullet point).*

Question 45: Is there a way to get a better description or methodology of the "loading" process? I'm not sure I understand it.

Measures that are more consistent, or more correlated, with other measures within the group have a greater influence on the hospital's group score. The influence of an individual measure on the group score is represented by the measure's "loading."

A loading is empirically derived for each measure in a group when applying the latent variable model (LVM); these statistically estimated measure loadings are regression coefficients based on maximum likelihood methods using observed data and are not subjectively assigned. A loading reflects the degree of the measure's influence on the group score relative to the other measures included in the same group. A measure's loading is the same across all hospitals. Measures with higher loadings are more strongly associated with the group score and the other measures within that group. All measures included in the Star Ratings have an effect on the group score; however, measures with higher loadings have a greater association (or impact) on the group score than measures with substantially lower loadings. The loadings for the July 2016 Star Ratings are reported in Appendix C.

Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

Question 46: Where can we find the loading values for our July 2016 Preview reports?

The loadings for the July 2016 Star Ratings are reported in Appendix C in the July 2016 Quarterly Update Specifications report on QualityNet.

Question 47: The Load factor may direct hospitals to focus on certain measures. If this is not the intended purpose, does it place a facility in a difficult position given the prominence star ratings may have?

Measures that are more consistent, or more correlated, with other measures within the group have a greater influence on the hospital's group score. The influence of an individual measure on the group score is represented by the measure's "loading."

A loading is empirically derived for each measure in a group when applying the LVM; these statistically estimated measure loadings are regression coefficients based on maximum likelihood methods using observed data and are not subjectively assigned. A loading reflects the degree of the measure's influence on the group score relative to the other measures included in the same group. A measure's loading is the same across all hospitals. Measures with higher loadings are more strongly associated with the group score and the other measures within that group. All measures included in the Star Ratings have an effect on the group score; however, measures with higher loadings have a greater association (or impact) on the group score than measures with substantially lower loadings. The loadings for the July 2016 Star Ratings are reported in Appendix C.

Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

Question 48: What is the implication of having a negative loading on model validity? Also, what does a negative loading mean?

A negative loading means that for a given quarter, performance on that measure was inversely related to the other measures in the group. However, the negatively loaded measures in July 2016 are likely to have a confidence interval that includes zero. The negative loading should not be over

interpreted.

MEASURES

Question 49: **How often will measures be added/removed? How are these changes being shared with hospitals and other key stakeholders?**

When measures are added to Hospital Compare, *they will be included in the star rating calculation if they meet all the criteria, that is, if they're not a Structural Measure. Any measure that's a measure of quality will be included in the star ratings.*

If a measure is retired or removed from the Inpatient or Outpatient Quality Reporting Programs, they will be removed from the star calculation, as well. The star rating is meant to be a summary depiction of the measure data on Hospital Compare.

Question 50: **With known (and published) impact of Socio-demographic status on many of these measures, including but not limited to readmission and HCAHPS, will you please exclude these from the calculation? This proven bias is contrary to the stated objective of providing consumers with a reliable, accurate, and simplified way to assess quality in a single score.**

CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve. In order to specifically address the issue of risk adjustment for socio-demographic status, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on this issue, as directed by the IMPACT Act; and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by the ASPE and consider if or how they apply to CMS quality measures and the Star Ratings.

Question 51: **How soon will NHSN CLABSI/CAUTI measures be expanded to include non-ICU patients?**

The CLABSI + selected wards (HAI-1) and CAUTI + selected wards (HAI-2) measures were included in the April and July 2016 Overall Hospital Quality

Star Rating results.

Question 52: **Many of the Effectiveness of Care Measures are no longer listed as Core Measures. How will these be reported?**

The Overall Hospital Quality Star Rating is a summary of measures reported on Hospital Compare *and does not publicly report individual measures.*

Question 53: **The two colonoscopy measures are new measures and have been reported only once. Typically, CMS collects a new measure for more than one year before inclusion in a program. What performance period will these measures be included in the Overall star rating?**

The Appropriate Follow-up Interval for Normal Colonoscopy in Average Risk Patients (OP-29) and Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use (OP-30) *measures added to the Process: Effectiveness of Care measure group have a data collection period of October 1, 2014 – September 30, 2015 for July 2016.*

Question 54: **Can you comment on the relationship between reduction of 30-day mortality rates and increased hospital-wide all-cause, unplanned readmission rate (HWR)? How do the Star Ratings account for this relationship? How would you explain this to your consumer groups? The literature has been very clear on the impact of socio-economic status (SES) and increased HWR. This one measure has many components unrelated to hospital quality of care. How will you address this moving forward?**

The HWR measure is endorsed by The National Quality Forum (NQF) and has undergone rigorous testing for scientific acceptability and validity. CMS believes that this measure is an important indicator of overall hospital quality.

CMS developed the Star Ratings to be as inclusive of as many measures currently reported on Hospital Compare *as possible in order to present the most comprehensive picture of hospital quality for consumers. The Star Ratings methodology includes a systematic process for determining the eligibility of a measure for inclusion which was vetted by the TEP and public comment.*

CMS uses the following criteria to exclude measures from the Star Ratings calculation:

- *Measures suspended, retired, or delayed from public reporting on Hospital Compare;*
- *Measures with no more than 100 hospitals reporting performance;*
- *Structural measures;*
- *Measures for which it is unclear whether a higher or lower score is better (non-directional);*
- *Measures no longer required for the Inpatient Quality Reporting (IQR) Program or Outpatient Quality Reporting (OQR) Program; and*
- *Duplicative measures (e.g., individual measures that make up a composite measure that is also reported; or measures that are identical to another measure).*

In addition, by studying the star ratings data, CMS concluded that it is unlikely that any one measure precludes a given type of hospital from performing well. For example, a hospital that has poor performance on a single Safety of Care Measure, such as PSI-90, may still receive a high Safety of Care group score and a high star rating if that hospital performs well on the other included Safety of Care measures. Similarly, CMS does not believe that the removal of the PSI-90 and Hospital-Wide Readmission measures would materially change a hospital's Star Ratings results.

CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures, while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve. In order to specifically address the issue of risk adjustment for socio-demographic status, the ASPE is conducting research on this issue, as directed by the IMPACT Act; and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by ASPE and consider if or how they apply to CMS quality measures and the Star Ratings.

Question 55: Why is Home Management Plan of Care (CAC-3) included if CMS specifically excludes measures no longer required for IQR/OQR?

The CAC-3 measure was included in the July 2016 Overall Hospital Quality Star Rating as it is scheduled to be retired from the Hospital IQR Program with the anticipated December 2016 Hospital Compare release.

Question 56: **Many of the measures chosen for the star ratings essentially substitute the national average for the hospital's own performance to the extent the hospital has a small number of relevant cases. By using these measures, CMS has made it hard for high performing smaller hospitals to show that they have excellent performance, and yes the reverse is true, as well, that low performing small hospitals may appear better in the star ratings than they otherwise would look. How do you intend to explain to the public that these ratings do not really reflect the performance of small hospitals?**

The Overall Hospital Quality Star Ratings represent a performance summary designed to facilitate patient and consumer use of Hospital Compare. *The Overall Hospital Quality Star Ratings allow consumers to compare hospitals with greater understanding by simplifying detailed information for patients and conveying information on multiple dimensions of hospital quality in a single score. This effort responds to sections of the Affordable Care Act, which call for public reporting that is transparent, efficient, easily understood, and widely available. In addition, the Overall Hospital Quality Star Ratings serve to improve accessibility to* Hospital Compare *by giving hospitals an initial summary glance that allows them to further explore hospital quality through individual measures.*

CMS designed several aspects of the Overall Hospital Quality Star Ratings development process to include the patient and consumer perspective in key methodological, display and policy decisions. Both the TEP and patient and patient advocate working group included diverse patient and patient advocate representation. These individuals were supportive of CMS' decision to develop a hospital quality star ratings system, expressing its potential value and importance to patients and consumers. CMS will do continued consumer testing and patient workgroup engagement to further improve the display of information to reduce confusion and guide consumers.

CMS will develop resources to facilitate patients' and consumers' understanding of the methodology.

In addition, CMS will continue to assess the impact of the Overall Hospital Quality Star Ratings methodology on different types of hospitals to inform future improvements.

METHODOLOGY

Question 57: **Are the HCAHPS Experience measures based on the Top Box or the Linear Mean. Both are reported on *Hospital Compare*.**

In the case of the star ratings methodology for the HCAHPS measures, the scores that are used are the linear mean scores.

Question 58: **How did you determine the weighting of the measures? For instance, it may not be intuitive (or reasonable) that an outcome, such as Mortality, should carry equal weight to Readmissions.**

The methodology development process around the weighting involves multiple forms of feedback. When we initially developed the weights, we tried to use places where there was already some policy guidance, so there are weights that are used in the Hospital VBP Program that emphasize outcomes over process and equally emphasize different domains of outcomes. We took those weights and we vetted with them the multi-stakeholder TEP. We showed it to a patient advocate workgroup, and then we also had a public comment on them in the spring of last year.

We received a lot of support for the weighting and how the outcomes are emphasized over process, and how each of the outcomes groups were equally weighted. There's no gold standard, no correct or right number for weighting, but this seems to be consistent with a variety of policy programs and a place where there was initial consensus. We are very open to additional feedback or additional concerns from a methodology development perspective.

Question 59: **With the LVMs, what evidence can be provided about the extent to which a single latent dimension accounts for common variance among targeted measures for each LVM?**

Early in the development process when we sought to evaluate whether or not the Latent Variable Modeling approach would be appropriate, as well as meet the objectives of the star ratings, we did several factor analyses. What we found in our factor analysis was that the use of these measure groupings, as they are currently used, identifies one meaningful Latent Variable per group. this indicates that the assumption that the mortality measures together, for example, all reflect one common Latent Variable with respect to

mortality performance was fairly strong and robust.

The one exception to this was the measures in the Imaging Efficiency group where there may have been one to two predominant Latent Variables. That group is weighted very little towards the overall star rating, and when we vetted it with both our multi-stakeholder TEP, as well as the public through public comment period, the general consensus was that the principle of inclusiveness of measures meant that we should still include those measures and include that group as a distinct group.

Question 60: **K-means clustering is typically used when classifying objects based on multiple measures. In this case, k-means was applied to a single summary score. How does this effectively differ from breaking the summary score distribution into five ranked categories?**

The k-Means Clustering approach was one that we considered in the methodology development process alongside a few other options. We acknowledge that k-Means is often used to classify or cluster variables when you've got multiple vectors, but it can still work well with a single dimension variable or one variable, in this case the Hospital Summary Score.

We originally developed the methodology by considering several ways to classify or cluster hospitals into each of the star categories. The simplest method we devised was to draw five lines, or essentially classify hospitals into quintiles. Every hospital from zero to 19.999 would be the first quintile, and that would be one star; from 20 to 39.999 would the second quintile, and that would be two stars.

Another approach we considered was to set statistical thresholds and say that a hospital had to meet a rule, for example, a Hospital Summary Score that was statistically higher than the national average score and also greater than 50% of individual measures greater than the national average.

The third approach we considered, which is what we ultimately used, was the k-Means Clustering approach.

We took all of these approaches to the TEP, as well as a discussion in public comment, and the general consensus was that the k-Means Clustering would allow us to meet a variety of goals in that classification that the others didn't meet. For one thing, it didn't create an arbitrary line between percentiles (e.g., the 19th and 20th percentiles). Those hospitals may have a score that is nearly identical but they'd be getting a different star rating simply because

the line was made at a fixed point.

k-Means Clustering allows for the size of each group to be unequal. It allows for there to be a non-normal distribution of star ratings. In the case of this most recent reporting period, there are more three-star hospitals because many hospitals perform near the national average overall and across many of the measures that they report.

The other advantage of k-Means Clustering is that it intuitively groups hospitals together that have a hospital summary score that is more similar. So, when we've done subsequent testing of things such as the validity or using simulation to test the reliability of our classification of hospitals, we find that k-Means Clustering would perform better than other approaches, such as the quintiles approach.

This is a place where CMS has also said in the past that they're open to additional feedback and comments. It's likely a part of the methodology that will continue to evolve over time, but we think that this initial point is a good place to start in terms of classifying hospitals into these five groups.

Question 61: **The star rating methodology penalizes for single metrics that are "no different from the National average;" how is the general public, who needs a star rating, going to differentiate what is within the confidence intervals of the mean?**

CMS is exploring opportunities, both within the display and the support materials, to present concepts such as the confidence intervals (when relevant) to patients and consumers in a fashion that does not increase confusion but conveys the information within the star ratings.

Question 62: **Where can we see the star ratings calculations if the Latent Variable Modeling method was not used? In other words, what is the effective impact of this adjustment?**

CMS employs LVM to estimate a group score for the dimension of quality represented by the measures in each group. LVM is a statistical modeling approach that assumes each measure reflects information about an underlying, unobserved dimension of quality. A separate LVM is constructed for each group so that a total of seven latent variable models are used to calculate the Star Ratings. The LVM accounts for the relationship, or correlation, between measures for a single hospital. Measures that are more

consistent with each other, as well as measures with larger denominators, have a greater influence on the derived latent variable. The model estimates for each hospital the value of a single latent variable representing the underlying, unobserved dimension of quality; this estimate is the hospital's group score.

CMS chose this modeling approach based on statistical literature regarding aggregating healthcare quality measures and the previous use of the approach in other disciplines, such as psychology and education.

As noted above, measures that are more consistent, or more correlated, with other measures within the group have a greater influence on the hospital's group score. The influence of an individual measure on the group score is represented by the measure's "loading."

A loading is produced for each measure in a group when applying the LVM; these statistically estimated measure loadings are regression coefficients based on maximum likelihood methods using observed data and are not subjectively assigned. A loading reflects the degree of the measure's influence on the group score relative to the other measures included in the same group. A measure's loading is the same across all hospitals. Measures with higher loadings are more strongly associated with the group score and the other measures within that group. All measures included in the Star Ratings have an effect on the group score; however, measures with higher loadings have a greater association (or impact) on the group score than measures with substantially lower loadings. While empirically calculated loadings may not match conceptual frameworks of measure importance, CMS believes the strengths of this approach outweigh this limitation.

Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

Please refer to the "Comprehensive Methodology Report (v2.0)" and "Quarterly Update and Specifications Report (v2.2) (July 2016)" posted on QualityNet Star Ratings Page for more information regarding the Star Rating methodology.

Question 63: **On slide 30, VTE 1, VTE 2, and VTE 3 were already retired. So how does CMS have these data to compare?**

VTE-1, VTE-2, and VTE-3 are scheduled to be retired from the Hospital IQR program with the December 2016 Hospital Compare *release.*

Question 64: **The LVM assumes that the commonality you are seeing in the measure is related to the underlying quality, but it may instead be related to the complexity of the population served or to socio-demographic factors that are not accounted for in the measures, but are affecting outcomes for patients. Or, they may be related to other factors. What testing have you done to see if "quality" is truly the latent factor that is being measures?**

Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

Early in the development process when we sought to evaluate whether or not the LVM approach would be appropriate, as well as meet the objectives of the star ratings, we did several factor analyses, and what we found in our factor analysis was that the use of these measure groupings, as they're currently used, identifies one meaningful Latent Variable per group. This supports the assumption that the mortality measures together, for example, all reflect one common Latent Variable with respect to mortality performance as strong and robust.

The one exception to this was the measures in the Imaging Efficiency group where there may have been one to two predominant Latent Variables. That group is weighted very little towards the overall star rating, and when we vetted it with both our multi-stakeholder TEP, as well as the public, through a public comment period, the general consensus was that the principle of inclusiveness of measures meant that we should still include those measures and include that group as a distinct group.

Question 65: **When developing the group scores in the latent variable model, were confidence intervals for the measure scores taken into account? We noticed on one of our group scores, we were listed as "Worse than the National Rate," while at the measure level, all of the measures included in that group were "No Different than the National Average." Could you explain more about how this could occur?**

CMS does not calculate the Group Scores for the Overall Hospital Quality

Star Ratings as a simple average of individual measure scores. Before calculating group scores, the star rating methodology standardizes all included measures and ensures they are in the same direction (i.e., a higher score indicates better quality). This step distinguishes the star rating group scores from individual measure scores, especially Mortality and Readmission measure scores. Each measure group score, or point estimate, has an associated variance in the form of a 95% confidence interval. There are many factors that can influence the width of the confidence interval for a hospital's measure group score. One of those factors is the number of measures a hospital reports within that measure group. For example, a hospital that reports more measures in a measure group may have a narrower confidence interval than a hospital with fewer measures. In addition, hospitals with larger denominators for a given measure are more likely to have a more precise score for each individual measure, which may also result in a narrower 95% confidence interval. Next, group scores are calculated using LVMs. The model assigns a "loading" to each measure in the group. The measure loading, empirically derived and consistent across all hospitals, quantifies a measure's impact on the group score. In other words, if a measure has a higher loading, the measure may have a greater impact on the group score than measures with lower loadings. For each measure group, the confidence interval of a hospital's group score is compared to zero to assign a national comparison category according to the following guidelines:

- *If the hospital's interval falls entirely above zero, the score falls "Above the national average"*
- *If the hospital's interval includes zero, the score is the "Same as the national average"*
- *If the hospital's interval falls entirely below zero, the score falls "Below the national average"*

The measure group score does not directly translate into a national performance category since the 95% confidence interval is required to compare the measure group score to the national average.

Question 66: **Is CMS planning on addressing the concerns voiced by the Association of American Medical Colleges (AAMC) and other groups across the country about the distribution of scoring for academic facilities vs. community facilities? As many have stated, this methodology does not account for the patients cared for in an academic setting vs. community setting.**

The Overall Hospital Quality Star Ratings represent a performance summary based on specific measures available on Hospital Compare. *The methodology seeks to be inclusive of as many measures and hospitals as possible, regardless of characteristics or size. In doing so, the statistical model includes all available information and uses standard inclusion criteria and calculations for all included hospitals.*

We understand the challenges some academic medical centers may face due to the complexities of their patients. While the Star Rating methodology does not risk adjust for patients transferred from a smaller hospital to an academic medical center, several of the underlying measures do exclude these patients. For example, the Condition-Specific Mortality Measures hold the transferring facility accountable, not the receiving facility.

CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.

Question 67: Are there any plans to adjust for teaching hospitals vs non-teaching hospitals in the future?

The Overall Hospital Quality Star Ratings represent a summary of performance based on specific measures currently available on Hospital Compare.

Responsiveness to stakeholder feedback is a guiding principle for the development and future refinements of the Star Ratings methodology. CMS has made substantial efforts to engage with the public and hospitals on the Star Ratings, including two (2) public comment periods, TEP meetings, and a national stakeholder call. CMS was purposeful in ensuring that the hospital perspective was represented on the TEP, which includes several nominees from the AAMC, American Hospital Association (AHA), and state hospital associations. In addition, CMS has provided ongoing support to hospitals, responding to their individual questions during the July 2015 hospital dry run and April 2016 Preview Period.

CMS will continue to examine the impact of the Overall Hospital Star Ratings methodology on hospitals to inform future improvements.

CMS will continue this engagement with stakeholders through continued consultation of the TEP, possible public comment periods, Hospital Compare

support calls and materials, National Provider Calls, and active solicitation of feedback to inform future improvements to the methodology.

Question 68: **I have concern that the CMS methodology does not recognize hospitals that serve large populations of trauma and oncology patients, as well as a disproportionate portion of medically underserved patients. How will CMS address this flaw in the methodology?**

CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve. In order to specifically address the issue of risk adjustment for socio-demographic status, the ASPE is conducting research on this issue as directed by the IMPACT Act; and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by ASPE and consider if or how they apply to CMS quality measures and the Star Ratings.

Question 69: **In the star rating determination, is there any stratification for the size of the hospital and the level of services offered, as these factors can impact a facilities quality measures outcomes?**

The hospital size and level of services offered are currently not used to determine a hospital's star rating. CMS will continue to assess the impact of the Overall Hospital Quality Star Ratings methodology on different types of hospitals to inform future improvements.

Question 70: **Was there one measure specifically that had a big sway over the rating?**

CMS developed the Star Ratings to be as inclusive of as many measures currently reported on Hospital Compare as possible in order to present the most comprehensive picture of hospital quality for consumers. The Star Ratings methodology includes a systematic process for determining the eligibility of a measure for inclusion which was vetted by the TEP and public comment.

CMS uses the following criteria to exclude measures from the Star Ratings calculation:

- *Measures suspended, retired, or delayed from public reporting on Hospital Compare;*

- *Measures with no more than 100 hospitals reporting performance;*
- *Structural measures;*
- *Measures for which it is unclear whether a higher or lower score is better (non-directional);*
- *Measures no longer required for the IQR or OQR Programs; and*
- *Duplicative measures (e.g., individual measures that make up a composite measure that is also reported; or measures that are identical to another measure).*

In addition, by studying the star ratings data, CMS concluded that it is unlikely that any one measure precludes a given type of hospital from performing well. For example, a hospital that has poor performance on a single Safety of Care measure, such as PSI-90, may still receive a high Safety of Care group score and a high star rating if that hospital performs well on the other included Safety of Care measures. Similarly, CMS does not believe that the removal of the PSI-90 and Hospital-Wide Readmission measures would materially change a hospital's Star Ratings results.

Question 71: **United Healthcare is planning to release its own hospital quality star rating report to the public based on the 2013 CMS Medicare Provider Analysis and Review (MEDPAR) data file; their report and methodology contradict CMS' methodology and ratings. Have you seen their report and methodology?**

CMS has not reviewed the methodology used by United Healthcare. However, the Overall Hospital Quality Star Rating methodology is designed to be inclusive of the data publicly reported on Hospital Compare. *Other hospital ratings systems may be based on different measures and data sources. Other hospital quality star ratings efforts,* including Nursing Home Compare, Dialysis Facility Compare *and* Home Health Compare, *utilize different methodologies for the development of their star ratings due to the differences in types of measures and variances in the amount of information available.*

Question 72: **Does the weighted likelihood method used to develop the group scores in the LVM imply that larger hospitals, which tend to constitute urban or academic centers, have more weight in developing the loadings than other, smaller, more rural facilities? If so, what type of testing was completed to show that these larger hospitals did not have statistically different scores from smaller hospitals?**

Hospitals' reported measure scores may include different numbers of patients, depending on the measure. For each measure, some hospitals may report a score based on data from fewer cases while other hospitals report scores based on more cases, resulting in differing precision for each hospital's individual measure score. This variability in precision is usually known as "sampling variation."

CMS gives more weight to measure scores that are more precise by using a Weighted Likelihood Method. This method uses the hospital's measure denominator (hospital case count or sample size) to weight the observed value. A weighted likelihood ensures that a hospital with a larger denominator, or a more precise measure score, contributes more in calculating the loadings used to estimate the group score.

Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.

MISCELLANEOUS

Question 73: **Will CMS share the calculation workbook so we can supply our most recent data into the workbook to anticipate what the Star Rating will be? (The Star Rating CMS publishes usually uses "older data.")**

We are not 100 percent sure to what workbook you are referring. CMS is exploring the best way to release the statistical software package and the code used to calculate the star ratings. We are exploring the feasibility of doing this. When we decide to do so, we will send notification to all hospitals so that they can run their own data.

Question 74: **Question regarding distribution of stars: were academic medical centers skewed toward the lower end?**

The Overall Hospital Quality Star Ratings represent a performance summary based on specific measures available on Hospital Compare. *The methodology seeks to be inclusive of as many measures and hospitals as possible, regardless of characteristics or size. In doing so, the statistical model includes all available information and uses standard inclusion criteria and calculations for all included hospitals.*

We understand the challenges some academic medical centers may face due to the complexities of their patients. While the Star Rating methodology does not risk adjust for patients transferred from a smaller hospital to an academic medical center, several of the underlying measures do exclude these patients. For example, the condition specific mortality measures hold the transferring facility accountable, not the receiving facility.

CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.

Question 75: **For Patient Experience, the Hospital VBP program does not include the "Willingness to Recommend Hospital" measure in their program. Why is it included here?**

The overall star ratings include all 11 measures used in the HCAHPS star ratings to be consistent. Please note that stars are not used for payment.

Question 76: **We are concerned that billing "administrative data" continues to be used for measuring clinical outcomes. Is there a move to use clinical data, as we feel clinical "registry data" is more in alignment with clinical outcomes and reflective of clinical care provided by a hospital?**

CMS is exploring the development of hybrid measures, obtaining some data from administrative claims and clinical factors from the EHR. These measures are in the early stages of development, and implementation will be determined through rulemaking with a comment period.

Question 77: **How many hospitals had a score of 1?**

Please refer to July 2016 Quarterly Updates and Specifications report on Page 13, Table 4 located on QualityNet. *1 Star = 133.*

Question 78: **Will CMS be publishing the modeling diagnostics/performance so that the public and researchers can provide input and feedback for improvement?**

CMS provided a comment period in the summer of 2015 seeking feedback on

the methodology. There are no plans to seek further comment; however, we will take this into consideration if the methodology is updated in any way.

Question 79: **What are the reasons why CMS is not using Managed-Medicare patient population vs. using only Traditional Medicare Fee-for-Service patient population?**

CMS does not have access to Medicare Advantage Plan claims data, as those claims are received by participating private insurers.

Question 80: **Are we comfortable the acuity found in Level I Trauma centers is adequately adjusted for in the star methodology?**

Yes, we believe the measures, specifically the Outcomes Measures, are adequately risk adjusted to account for trauma cases. In many scenarios, they are excluded from the measures if they do not meet the condition-specific criteria.

Question 81: **Is there a way to drill down cases included in the reporting period if/when validation/focus study is needed?**

Data at the patient level is available in your IQR and OQR HSRs.

OP-22

Question 82: **Please define the Reporting Period for OP-22 that will be updated in July 2016. The Preview Report does not provide that information.**

Reporting period for OP-22 ED-Patient Left without Being Seen: 1/1 – 12/31/14 (1Q14-4Q14)

OP-29 and OP-30

Question 83: **I thought no measure could be used unless it was up on *Hospital Compare* for a year. How can the colonoscopy measures, OP-29 and OP-30, be included in the Effectiveness of Care Measures?**

The methodology is designed to be as inclusive as possible for measures used on Hospital Compare.

CMS uses the following criteria to exclude measures from the Star Ratings calculation:

- *Measures suspended, retired, or delayed from public reporting on Hospital Compare;*
- *Measures with no more than 100 hospitals reporting performance publicly;*
- *Structural measures;*
- *Measures for which it is unclear whether a higher or lower score is better (non-directional);*
- *Measures no longer required for IQR or OQR Programs; and*
- *Duplicative measures (e.g., individual measures that make up a composite measure that is also reported; or measures that are identical to another measure).*

Because OP-29 and OP-30 do not meet any of these exclusion criteria, they are included in Star Ratings calculation for July.

OUTCOME READMISSION

Question 84: **Our hospital's Overall Hospital Star Rating Outcome: Readmission group score is below the National Average when the 30-day risk-standardized condition-specific and hospital-wide readmission rates are no different than the National Rate. Why is so?**

The easiest way to think about this is to remember that individual measures when they are reported will have a different national average score and a different 95 percent confidence interval that needs to be exceeded or be below of in order to be called above or below the national average.

In the case of a group score, we start to combine summarized information across many measures. If a hospital does moderately well at one readmission measure, it may be possible that it didn't meet the threshold to be above the national average for that individual measure, but if they do well across all of the readmission measures, in some in combination, that might be very high performance in comparison to other hospitals; therefore they may end up with a group score that's above the national average.

This is one of the features of the star ratings performance categories that

seeks to show and summarize performance that is strong or poor on the spectrum.

Question 85: For Outcome Readmission, is Above National Average a positive or negative outcome?

The easiest way to think about this is to remember that individual measures when they are reported will have a different National Average Score and a different 95 percent Confidence Interval that needs to be exceeded or be below in order to be called above or below the national average.

*In the case of a group score, we start to combine summarized information across many measures. If a hospital does moderately well at one readmission measure, it may be possible that it didn't meet the threshold to be above the national average for that individual measure, but if they do well across **all** of the readmission measures, in some in combination, that might be very high performance in comparison to other hospitals; therefore, they may end up with a group score that's above the national average.*

This is one of the features of the star ratings performance categories that seeks to show and summarize performance that is strong or poor on the spectrum.

OVERALL SUMMARY SCORE

Question 86: Please explain the overall summary score in the preview report?

CMS calculates your hospital's Overall Summary Score as a weighted average of your available group scores; each group score is weighted according to the weighting scheme reflected in your preview report. If your hospital does not have all measures in all measure groups, then measure group weights are re-proportioned to sum to 100%.

CMS applies k-Means Clustering to the national sample of overall summary scores in order to sort hospitals into one of five star categories.

PEER DATA

Question 87: Where can we find peer data (bed size, teaching status) on the group scores and overall star ratings?

This information is currently not provided for the Overall Star Rating. CMS will take this into consideration for future releases.

Question 88: **Our hospital administration would be interested in knowing if there are any large acute care hospitals who have achieved 5-star ratings in patient experience.**

For Hospital Compare Overall Hospital Star Ratings, no all-inclusive list is publicly available. CMS anticipates publicly reporting the Overall Star Ratings in a future release yet to be determined.

PERFORMANCE PERIOD

Question 89: **What is the performance period for this data?**

The reporting periods for each measure group included in the Overall Hospital Quality Star Rating are outlined below:

- *Safety*
  - *HAI measures: 10/1/14 – 9/30/15 (4Q14–3Q15) except HAI-1 and HAI 2: 1/1/15 – 9/30/15 (1Q15–3Q15)*
  - *COMP-Hip-Knee: 4/1/12 – 3/31/15 (2Q12–1Q15)*
  - *PSI 90: 7/1/13 – 6/30/15 (3Q13–2Q15)*
- *Mortality*
  - *MORT measures: 7/1/12 – 6/30/15 (3Q12–2Q15)*
  - *PSI-4 measure: 7/1/13 – 6/30/15 (3Q14–2Q15)*
- *Readmissions: 7/1/12 – 6/30/15 (3Q12–2Q15)*
- *Patient Experience: 10/1/14 – 9/30/15 (4Q14–3Q15)*
- *Effectiveness of Care: 7/1/14 – 6/31/15 (3Q14–2Q15)*
  - *IMM-2 Influenza Immunization: 10/1/14 – 3/31/15 (4Q14–1Q15)*
  - *IMM-3/OP-27 Healthcare Personnel Influenza Vaccination: 10/1/14 – 3/31/15 (4Q14–1Q15)*
  - *OP-22 ED-Patient Left without Being Seen: 1/1 – 12/31/14 (1Q14–4Q14)*
  - *OP-29 Endoscopy/Polyp Surveillance: Appropriate Follow-Up Interval for Normal Colonoscopy: 1/1/14 – 12/31/14*
  - *OP-30 Endoscopy/Polyp Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps: 1/1 – 12/31/14*
- *Timeliness of Care: 10/1/14 – 9/31/15 (4Q14–3Q15)*
- *Efficient Use of Medical Imaging: 7/1/14 – 6/30/15 (3Q14–2Q15)*

PHONE CONTACT

Question 90: **Why is there not a phone number that you can call to get help[regarding the Star Ratings Program]? I tried to get an answer through email but it has been very difficult.**

Unfortunately, due to multiple contractors and/or subject matter experts being involved in the Overall Hospital Quality Star Ratings program, it is difficult to coordinate a phone call.

PREVIEW REPORT

Question 91: **Where do we download the preview reports on QNet? Are these for IQR only or combined with OQR preview reports?**

The Hospital IQR and OQR July 2016 preview reports can be accessed via the QualityNet Secure Portal *by:*

- *Navigating to the public website for* QualityNet *at* https://www.qualitynet.org
- *Selecting* ***[Login]*** *under the "Log in to* QualityNet Secure Portal*" header*
- *Entering your* QualityNet *User ID, Password, and Security Code and selecting* ***[Submit]***
- *Reading the Terms and Conditions statement and selecting* ***[I Accept]*** *to proceed*

Preview reports can be run by:

- *Selecting* ***[Run Reports]*** *from the "My Reports" drop-down*
- *Selecting* ***[IQR" or "OQR]*** *from the "Report Program" drop-down*
- *Selecting* ***[Public Reporting – Preview Reports]*** *from the list in the "Report Category" drop-down*
- *Selecting* ***[View Reports]****; the selected report will display under "Report Name"*
- *Selecting* ***[Public Reporting – Preview Reports]*** *under "Report Name"*
- *Selecting* ***[Run Reports]***
- *Select the* ***[Search Reports]*** *tab.*

The report requested will display, as well as the report status. A green check mark will display in the "Status" column when the report is complete. Once

complete, the report can be viewed or downloaded.

Question 92: What would be the best method/or process to take to determine why a star rating may have changed?

We suggest looking and comparing your preview reports from the previous release and current release and then taking into consideration performance on any of the new measures that may have been added, as well as considerations of performance of any measures that may have been removed from the programs, and then ultimately from the website.

Question 93: Which quarter do the results in the most recent *Hospital Compare* preview report (outpatient) represent?

The July 2016 Inpatient/Outpatient Preview report released May 6 is for July release.

Question 94: I never got a report with the group ratings on it. How would I go about getting this again with April and July on it? Where do we download the preview reports on QNet? Are these for IQR only or combined with OQR preview reports?

The Hospital IQR and OQR July 2016 preview reports can be accessed via the QualityNet Secure Portal by:

- *Navigating to the public website for QualityNet at https://www.qualitynet.org*
- *Selecting [Login] under the "Log in to* QualityNet Secure Portal*" header*
- *Entering your QualityNet User ID, Password, and Security Code and selecting [Submit]*
- *Reading the Terms and Conditions statement and selecting [I Accept] to proceed*

Preview reports can be run by:

- *Selecting "Run Reports" from the "My Reports" drop-down*
- *Selecting "IQR" or "OQR" from the "Report Program" drop-down*
- *Selecting "Public Reporting – Preview Reports" from the list in the "Report Category" drop-down*
- *Selecting "View Reports"; the selected report will display under*

"Report Name"
- *Selecting "Public Reporting – Preview Reports" under "Report Name"*
- *Selecting [Run Reports]*

Select the [Search Reports] tab. The report requested will display, as well as the report status. A green check mark will display in the "Status" column when the report is complete. Once complete, the report can be viewed or downloaded

Question 95: **How did our mortality group score go from 0.34 in the April 2016 preview report to -0.36 in the July preview report? It doesn't seem likely that our mortality performance changed that drastically in a single quarter.**

We suggest looking and comparing your preview reports from the previous release and current release and then taking into consideration performance on any of the new measures that may have been added, as well as considerations of performance of any measures that may have been removed from the programs, and then ultimately from the website.

Question 96: **Why are preview reports only made available through QNet for a certain time frame?**

Via rulemaking and comment period, CMS determined the preview period of 30-days was adequate for hospitals to be able to download and store their own reports. With limited infrastructure for archiving preview reports, CMS only allows 30-days for the preview period.

PSI-90

Question 97: **AHRQ submitted a revised measure set for the PSI-90 for approval to NQF. In its submission, AHRQ identified several significant biases and flaws in the current methodology. This information became available after the advisory panel meetings occurred. Can you explain why CMS has not eliminated this variable from the Star Rating methodology, given the known limitations and impact on the Star Rating?**

At this time, CMS is working with our star ratings team to do some analysis on the measure and determine the best way to move forward. The hospital

star ratings were designed to be as inclusive of as many measures that are currently publicly reported on the Hospital Compare *website. and that is one reason why the PSI-90 measure has not been removed from the star ratings.*

Question 98: **AHRQ PSI 90 is one element of the overall stars. AHRQ has already acknowledged that they will not be able to calculate an O: E ratio for at least one year, and probably two -- given the conversion from ICD-9 to ICD-10. Exactly how will the AHRQ PSI 90 metrics be incorporated into the overall star calculation as crude rates?**

The overall star ratings methodology takes the current publicly reported results and calculates the star ratings. As the individual measures and public reporting evolves, the star ratings methodology has been designed to adapt to how measures are publicly reported during that quarter.

Question 99: **When reviewing the April *Hospital Compare* preview reports and corresponding methodology report, we noticed that the empirically-calculated loadings for the Safety of Care group were significantly skewed to heavily weigh the PSI-90 Composite score compared to the other measures in the group (PSI-90 Loading = 0.92, all others were 0.29 or lower). Could you explain more about how this occurred, and if there is anything built into the model to prevent one measure from influencing the group score significantly?**

The influence of an individual measure on the group score is represented by the measure's "loading."

A loading is produced for each measure in a group when applying the LVM; these statistically estimated measure loadings are regression coefficients based on maximum likelihood methods using observed data and are not subjectively assigned. A loading reflects the degree of the measure's influence on the group score relative to the other measures included in the same group. Key considerations for measure loadings include:

- *A measure's loading is specific to the measure, considering national performance on the measure and the measure's relationship to other measures in the group and the group's latent variable. It is the same for all hospitals reporting that measure.*
- *Measures with higher loadings are more strongly associated with the group score. These more "consistent" measures, in terms of hospital performance, give us more signal or information about a hospital's*

quality profile than measures with "random" performance. Loadings are estimated using maximum likelihood. If several measures all point consistently in one direction, but one points in the opposite direction, the outlier receives less loading.

- *Large measure loadings do not directly imply that only a few measures "matter" towards the group score. However, measures with higher loadings do have a greater association (or 'impact') on the group score than measures with much lower loadings. There could be multiple measures with large loadings in one group. Measures that are reported by more hospitals with consistent performance will tend to have higher loadings, as they reflect a stronger "signal" of hospital quality.*
- *Given that CMS will re-estimate the loadings each time the Star Ratings are updated, the loadings for an individual measure can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.*

The PSI-90 measure is likely to have received the largest loading because performance on the measures is highly correlated with all other measures in the Safety of Care group. In addition, large sample or denominator sizes for many hospitals with PSI-90 scores are interpreted to indicate more precise measurements, which may also result in higher loadings.

Question 100: **PSI-04 is not a validated quality measure. This measure is so old it does not even accommodate "POA" (present on admission) = no. In other words, we are being 'charged' with poor quality when the patient arrives with the condition. Why is this measure being treated as a validated quality measure and included in the mortality calculation? PSI-04 should be retired because it has not been shown to be related to clinical quality. What is the evidence that this measure should be used for quality reporting?**

Please see the response to this frequently asked question here:

http://www.qualityindicators.ahrq.gov/FAQs_Support/FAQ_Software.aspx#sec04

Look for the question "PSI: What logic was used in the development of PSI #4 Death Rate among Surgical Inpatients with Serious Treatable Complications?"

References with empirical evidence are included in the response to the frequently asked question.

Please note that the intent of the QI measures is not to "charge" any particular facility. The AHRQ QI software calculates an observed rate that includes all incidents meeting PSI 04 criteria along with a risk adjusted rate that uses an expected PSI 04 rate based on the severity of illness in the population eligible for the measure. A facility with a higher severity of illness will have a lower risk adjusted result while a healthier population may have a higher one, allowing comparison with other facilities with a different population.

Comparison of rates among facilities is best done using the risk adjusted rate that controls for severity of illness, not the observed rate. POA is taken into account in the risk adjustment calculation for PSI 04.

Question 101: Will PSI 90 be removed for Maryland's hospitals given the delayed requirement for use of the POA indicator and ability of the contractor (Novitas) to accept corrected claims?

CMS will continue assess state based differences in measure reporting and the impact on star ratings. In order to ensure consistency with Hospital Compare *and minimize confusion for consumers and patients, the current star ratings are designed to include all measures currently reported on* Hospital Compare *in star ratings.*

Question 102: Will CMS consider the recent study published by Winters in Medical care that evaluated the validity of the PSIs in deciding whether to continue to report the PSI 90?

At this time, CMS is working with our star ratings team to do some analysis on the measure and determine the best way to move forward. The hospital star ratings were designed to be as inclusive of as many measures that are currently publicly reported on the Hospital Compare website, and that is one reason why the PSI-90 measure has not been removed from the star ratings.

PUBLIC COMMENT PERIOD

Question 103: Could CMS implement a public comment period to vet the inclusion of new measures?

The goal of the Star Ratings program is to be as inclusive of measures

currently publicly reported as possible. In the spring of 2015, CMS provided a 30-day comment period on the criteria for measure selection and exclusion from the methodology. The criteria are outlined in the methodology report which is publicly available on the QualityNet *Star Ratings page. At this time, there are no plans to seek comment on the inclusion of individual measures.*

PUBLIC REPORTING OF STAR RATINGS

Question 104: Will the actual star numbers for each hospital be available for download along with the data?

Yes, they are available on your Preview Report under the Overall Star Rating Section.

Question 105: When will Overall Star Ratings be reported on *Hospital Compare*?

CMS anticipates publicly reporting the Overall Star Ratings in a future Hospital Compare release to be determined.

The data used to generate Star Ratings must be from the same reporting data as the data on Hospital Compare *and the data from other hospitals used to generate the national average*

QUARTERLY REFRESH

Question 106: For the quarterly refresh, will you refresh the Mortality outcomes?

Measure loadings are updated on a quarterly basis. Refer to page 9 of the updated Methodology Report for the July 2016 Overall Star Rating posted on the QualityNet *website. Please note, the loadings for an individual measure are re-estimated each time the Star Ratings are updated and can dynamically change as the distribution of hospitals' performance on the measure and its correlation with other measures evolve over time.*

READMISSION

Question 107: Slide 27 says that 30-day Readmission will be considered also for THK. So, is a change being made to THK measures, as the PDF available considers complications until 90 days for some of the complications like

Mechanical Complication and Joint Infection?

Please submit this question to the measure developer at cmsreadmissionmeasures@yale.edu.

RISK-ADJUSTED

Question 108: Are these rates risk adjusted? Such as the HF or mortality rates?

The star rating in and of itself is not risk-adjusted. However, if the underlying measures are risk-adjusted, we didn't do anything to change those calculations.

STATISTICAL ANALYSIS SOFTWARE (SAS) PACK

Question 109: Is it possible for hospitals to replicate the star rating calculations in order to better monitor their ongoing performance?

CMS is exploring the feasibility of releasing the SAS packaging code and the national input file. (See slide #57 and #58.)

Question 110: What does SAS stand for?

Statistical Analysis Software

SCORING METHODOLOGY

Question 111: If a hospital's "Group score" is a positive score and higher than the National group score, is that considered to be "favorable?"

Yes, but the degree to which that score above average is categorized as "Above the National Average" depends on the confidence interval.

Question 112: Because our overall rating is based on previous data in *Hospital Compare*, how does being a new facility with no previous data affect our Star Rating and score?

If your hospital does not have at least three measure groups (one being an outcome group) with three or more measures, your facility will not receive a

star rating. If a hospital has too few measures, either due to changes in the measures available for Star Ratings or small volumes resulting in fewer individual measures in public reporting, the hospital star rating is replaced with "Not Available" accompanied by a footnote explaining that the hospital has too few cases to accurately calculate a star rating.

Question 113: How can our Average process times be less than State and National Averages yet we have a Negative Score?

CMS does not calculate the Group Scores for the Overall Hospital Quality Star Ratings as a simple average of individual measure scores. Before calculating group scores, the star rating methodology standardizes all included measures and ensures they are in the same direction (i.e., a higher score indicates better quality). This step distinguishes the star rating group scores from individual measure scores, especially Mortality and Readmission measure scores. Each measure group score, or point estimate, has an associated variance in the form of a 95% Confidence Interval. There are many factors that can influence the width of the confidence interval for a hospital's measure group score. One of those factors, is the number of measures a hospital reports within that measure group. For example, a hospital that reports more measures in a measure group may have a narrower confidence interval than a hospital with fewer measures. In addition, hospitals with larger denominators for a given measure are more likely to have a more precise score for each individual measure, which may also result in a narrower 95% Confidence Interval. Next, group scores are calculated using LVMs. The model assigns a "loading" to each measure in the group. The measure loading, empirically derived and consistent across all hospitals, quantifies a measure's impact on the group score. In other words, if a measure has a higher loading, the measure may have a greater impact on the group score than measures with lower loadings. For each measure group, the Confidence Interval of a hospital's group score is compared to zero to assign a national comparison category according to the following guidelines:

- *If the hospital's interval falls entirely above zero*
  - *"Above the national average" = better performance*
- *If the hospital's interval includes zero*
  - *"Same as the national average"*
- *If the hospital's interval falls entirely below zero*
  - *"Below the national average" = worse performance*
- *The measure group score does not directly translate into a national performance category since the 95% confidence interval is required*

to compare the measure group score to the national average.

SOCIO-DEMOGRAPHIC STATUS (SDS)

Question 114: **As a hospital system with both a medical center and rural setting hospitals, we are concerned that Medical Centers are at a disadvantage because of the increased number of patients, more measures included and 'sicker' patients who are more at risk for some of these complications. Likewise, rural setting entities are disadvantaged as we know zip code accounts for large percentage of variation in hospital readmissions.**

The Overall Hospital Quality Star Ratings represent a summary of performance based on specific measures currently available on Hospital Compare. *CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures, while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.*

In order to specifically address the issue of risk adjustment for SDS, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on this topic, as directed by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act); and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by the ASPE and consider if or how they apply to CMS quality measures and, ultimately, the Star Ratings.

Question 115: **I would like to know how CMS is going to address the issue of the lack of socio-demographic adjustments for the readmission and other outcome measures. This is particularly important for rural community hospitals.**

The Overall Hospital Quality Star Ratings represent a summary of performance based on specific measures currently available on Hospital Compare. *CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures, while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.*

In order to specifically address the issue of risk adjustment for SDS, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on this topic, as directed by the Improving Medicare

Post-Acute Care Transformation Act of 2014 (IMPACT Act); and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by the ASPE and consider if or how they apply to CMS quality measures and, ultimately, the Star Ratings.

Question 116: How do you consider hospitals that take higher risk patients or avoid influencing physicians to avoid high risk interventions to avoid poorer ratings?

The Overall Hospital Quality Star Ratings represent a summary of performance based on specific measures currently available on Hospital Compare. *CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures, while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.*

In order to specifically address the issue of risk adjustment for SDS, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on this topic, as directed by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act); and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by the ASPE and consider if or how they apply to CMS quality measures and, ultimately, the Star Ratings.

Question 117: Have you analyzed the impact of risk adjustment as a proportion of the inter hospital variation? Literature suggests that even Apache scoring does not adequately risk adjust patient severity for transfers. As a level 1 trauma, level 1 burn unit tertiary care referral community hospital, I am concerned that even small changes in numerators for mortality and PSI 4 will not be correctly adjusted by claims-based risk adjustment.

The overall star ratings take the measures that are currently on star ratings and only standardizes the scores to create the overall ratings. Recommendations or questions about individual measure methodology and risk adjustment should be sent to cmsstarratings@lantanagroup.com.

The mortality measures currently included are specific to Medicare FFS and include specific conditions such as AMI, heart failure, pneumonia, COPD, and Stroke.

Question 118: **Are nursing home readmissions, a known concern to CMS, in anyway adjusted for in the readmissions calculation?**

Nursing home readmission adjustment would require SDS risk adjustment.

The Overall Hospital Quality Star Ratings represent a summary of performance based on specific measures currently available on Hospital Compare. *CMS is committed to improving outcomes and working with stakeholders to improve individual quality measures, while minimizing unintended consequences for all facilities, regardless of the characteristics of the patients they serve.*

In order to specifically address the issue of risk adjustment for socio-demographic status (SDS), the Office of the ASPE is conducting research on this topic as directed by the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act); and will issue a report to Congress by October 2016. CMS will examine the recommendations issued by ASPE and consider if or how they apply to CMS quality measures and, ultimately, the Star Ratings.

SOFTWARE VERSION

Question 119: **Have you considered releasing a similar software package in R, not SAS, so that institutions will not have to buy an SAS license in order to understand their scores?**

In response to stakeholder feedback, CMS is considering publicly posting the Statistical Analysis Software (SAS) Pack with support documentation in the future for hospitals to have the ability to calculate their star rating. We are currently discussing the feasibility of providing the SAS code and national input file in a way that is transparent to each participating hospital, yet retains each hospitals privacy until the data is released nationally. Additionally, CMS is working to ensure that hospitals will have access to the complete Hospital Compare *downloadable dataset, including denominator data for each measure included in the Overall Star Ratings. It is important to keep in mind that there may be minor differences in the underlying scores due to software version and computer hardware.*

Question 120: **What AHRQ software version is currently being used to calculate the**

PSI-90?

The Hospital IQR Program used the fully recalibrated v5.0.1 AHRQ PSI software for Fiscal Year 2017.

STAR RATINGS

Question 121: Is the 5-Star rating limited to PPS hospitals or are Critical Access Hospitals included?

Any hospital that has data on Hospital Compare *and meets all of the inclusion criteria for measures, (has at least one of those outcome categories) can* **receive** *a star rating.*

Question 122: Will the star rating replace the current data listed for individual measures?

No, the Overall Hospital Quality Star Rating will not replace measures currently displayed on Hospital Compare. *The individual quality measure information used to calculate the Star Rating will continue to be displayed on* Hospital Compare*, and will be available for download in the database. The measures reported on* Hospital Compare *will continue to be updated based on regular rulemaking. The methodology for the Star Rating accommodates changes in the included measures over time (retirement of existing measures or addition of new measures).*

Question 123: Is there consideration of doing fractional stars as in some of the *Hospital Compare* programs?

Not at this time, but CMS will take this under consideration.

The overall rating on Hospital Compare *uses whole stars, which aligns with Star Ratings reported on other* Compare *websites, including* Nursing Home Compare *and* Dialysis Facility Compare. *Because these* Compare *websites and facility types are different, the Star Ratings for each site may vary in terms of the data on which they are based, how the stars are calculated, and how the stars are translated.*

Question 124: The number of required patient numbers and required number of reportable measures as restricted the ability of small critical access

hospital from achieving a star rating. Are there any future plans for a rating system for these small facilities?

The CMS team is currently analyzing the data we have for the April and July timeframes. If there are plans for a Star Rating program specifically for CAHs, it will be determined via rulemaking with comment period.

STAR RATINGS REFRESH

Question 125: **Is it reasonable to expect more star ranking movement when the measures that are only updated on an annual basis are scheduled to be updated?**

CMS plans to update the star ratings on Hospital Compare *and provide a preview report quarterly (April, July, October and December).*

Question 126: **Please clarify the star rating refresh for 2017 which will update semi-annually. Will this mean in April and October refresh the star rating will not be recalculated to match the measures that are updated every quarter?**

CMS plans to update the star ratings on Hospital Compare and provide a preview report quarterly (April, July, October and December).

SUMMARY SCORE

Question 127: **Slide 53: Can you provide more explanation about hospital re-calculate the summary score? For what purpose?**

CMS is considering making both the data file and SAS pack available to the public, including hospitals. This would allow hospitals to calculate not only their own scores, but the scores of other hospitals, as well.

TIMELINESS OF DATA

Question 128: **Why is there a nine-month time lag for the Press Ganey patient satisfaction data?**

The nine-month lag is a result of the behind-the-scenes submission timeline and processing that occurs prior to public reporting of the data. For

example, hospitals have up to 4.5 months after the close of the calendar quarter to finalize and submit their data to CMS.

TIMELINESS OF REPORTING

Question 129: **Many of the measures reported on *Hospital Compare* use different time frames. While this is understandable when comparing the same measure between hospitals, it seems concerning when we are combining quality measures from different time periods to represent one overall statistic to represent quality at a hospital at a specific point in time. For example, from the April preview report: Safety of Care – PSI-90s are 3rd Quarter 2012–2nd Quarter 2014; HAIs are from 3rd Quarter 2014–2nd Quarter 2015. Would you consider using newer, consistent time frames for all measures? Otherwise, public perceptions will be formed about a hospital's quality based on outdated information. If this will not be changed, will CMS consider adding a disclaimer underneath the Star Rating showing the time periods included in the calculation of the star measure?**

CMS uses the most-up-to date information available to us. The Claims-Based measures are calculated using the final paid claim, which minimizes significantly the chance a hospital will provide a revised claim, thereby changing the denominator. The Process of Care, HAI and HCAHPS data are submitted by the hospital to the various data warehouses. We give hospitals 4.5 months after the close of the quarter to submit their final, complete data file. The 4.5 months as agreed upon at the implementation of the IQR program, and was the timeframe requested by hospitals. After files are submitted, they are scrubbed, removing erroneous data, before the final measure calculations are made. The final calculations are then placed in the Hospital Compare *preview report which hospitals have 30 days to review. Once the preview period is over, any request from non-IPPS hospitals to suppress data, removal of closed hospitals, and any data issues are reconciled, and the final files for public reporting are posted.*

WINSORIZATION

Question 130: **Can you clarify the information on slides #21 and #18; they say different things about where the Winsorization happens. Which is right, or does it happen twice?**

*Winsorization is applied to hospital summary scores **prior to clustering**.*

CMS Winsorized the standardized measure score at the 0.125th percentile (Z= -3) and the 99.875 percentile (Z=3) of a Standard Normal distribution; thus, all standardized scores above 3 were set to be 3, and all standardized scores bellow -3 are set to be -3. This was done to avoid extreme outlier performance for which it is unclear if the reported measure score represented an extreme performance or potentially inaccurate reporting, as well as to avoid values that would make estimation technically challenging.

Question 131: Can you explain Winsorization more? I am not familiar with that.

Winsorization is used twice in the Overall Star Ratings methodology to reduce the effect of outliers.

First, we Winsorize individual measures scores. CMS utilizes Winsorization to limit the influence of measures with extreme outlier values at the 0.125th percentile (Z=-3) and the 99.875th percentile (Z=3). Winsorization is a common strategy used to set extreme outliers to a specified percentile of the data. All standardized measure scores above 3 are set to be 3, and all standardized below -3 are set to be -3. This has no material effect on the hospital group scores or star ratings, but make the computation more efficient.

Second, the methodology Winsorizes hospital summary scores prior to k-Means Clustering. The decision to Winsorize hospital summary scores, a modification from the Star Rating dry run, is based on comments received during the second public comment period and patients' and consumers' preference for a broader distribution of Star Ratings.

Question 132: For the Winsorization in Step 1 (of measures), do you Winsorize based on the actual percentile or rather based on the z-score being 3 standard deviations from the mean? If the latter, why is it appropriate to do a standard deviation based Winsorization when some measures are non-normally distributed, for example those with a floor of 0? The result of this method will be that some measures which have no Winsorization on one end (e.g., for infection measures, there might be no 3 standard deviations below the mean because that would be a negative number). Have you considered the impact of the Winsorization choice on scores?

The Winsorization is based on the z-score. CMS considered the impact of

Winsorization which resulted in modification of only 46 hospital summary scores. The decision to Winsorize hospital summary scores, a modification from the Star Rating dry run, is based on comments received during the second public comment period and patients' and consumers' preference for a broader distribution of Star Ratings.

CMS Winsorized the standardized measure score at 3 standard deviations from the mean of a Standard Normal distribution; thus, all standardized scores above 3 were set to be 3, and all standardized scores bellow -3 are set to be -3.

This was done to avoid extreme outlier performance for which it is unclear if the reported measure score represented an extreme performance or potentially inaccurate reporting, as well as to avoid values that would make estimation technically challenging. This Winsorization also improved the speed of computation without having any material impact on individual hospital group scores or star ratings.

END