NATIONAL QUALITY FORUM
Driving measurable health improvements together

Issue Brief

NQF Hospital Quality Star Rating Summit – 2020 Reconvening

Executive Summary

The Centers for Medicare and Medicaid Services (CMS) Overall Hospital Quality Star Ratings synthesize hospital quality information in a way that is easy for patients to understand. In September 2020, NQF convened a multistakeholder Expert Panel to provide input on several proposed updates to Star Ratings methodology for 2021 and subsequent years. The Panel discussed five select proposed changes and provided the following feedback to CMS.

- *Panel members were accepting of the proposed change to combine the three existing process measure groups, but highlighted several points for CMS to consider and test when selecting and grouping measures in the future.*
- *The Panel supported the proposed change to use a simple average of measure scores rather than the Latent Variable Model (LVM).*
- *The Panel supported the proposed update to the reporting threshold for receiving a Star Rating to ensure Star Ratings include measures in the priority areas of mortality or safety.*
- *There was a positive response to stratifying the Readmissions measure group by proportion of dual-eligible patients since it is aligned with the method used in the Hospital Readmissions Reduction Program (HRRP).*
- *There was not general agreement supporting the proposed approach to peer group hospitals by number of measure groups reported.*

Additional considerations for CMS to explore and test as part of ongoing efforts to refine the program are discussed further in this issue brief. This NQF-led convening reflects the value of multiple stakeholders collaborating to provide cohesive input on key components of quality measurement program and rating system methodology.

Background

On September 30, 2020, National Quality Forum (NQF) reconvened a multistakeholder panel of experts (Panel) from the inaugural Hospital Quality Star Rating Summit held on August 23, 2019 to discuss and provide feedback on several proposed changes to the CMS Overall Hospital Quality Star Ratings.

The recommendations from the August 2019 Summit served as a baseline for the Panel's discussion of the proposed changes. From the August 2019 convening, NQF made three recommendations to guide future actions for the Star Rating program.

1. Be clear about the program intent and goals. The program's methodology and design elements should align with the underlying intent and user needs.
2. Be transparent about what the Star Ratings do and do not convey. The Star Ratings are designed to provide a summary of measures on Hospital Compare to support consumer understanding. Clear communication about the interpretation of the program and its methods is imperative.
3. Design data presentation to meet consumer priorities and other user needs. There is an

opportunity to enhance the presentation of the Star Ratings and user interaction with the summary data to make the program more actionable and relevant.

The table below reflects the Panel's 2019 key considerations as they relate to several major proposed changes for the 2021 Star Ratings methodology. Note the table is meant to display alignment in topic area and provide examples of how changes match the Panel's considerations rather than suggest they are fully achieved.

No.	NQF Panel's 2019 Considerations	Select CMS Proposed Changes for 2021
1.	More closely align the construction of the program and design decisions with the underlying intent of the program	Increase simplicity of the methodology, predictability of measure emphasis within the methodology over time, and comparability of ratings among hospitals
2.	Consider regrouping measures to reflect clinically meaningful domains and service lines	Combine three existing process measure groups into one Timely and Effective Care group
3.	Expand the consideration of measures included in the program beyond Hospital Compare	
4.	Consider alternative, simpler approaches to group scoring that improve understandability and transparency of measure weighting	Use simple average methodology instead of LVM to calculate measure group scores
5.	Consider an explicit approach to determine Star Rating thresholds to enhance the predictability and actionability of the summary data	
6.	Consider eligibility criteria for hospitals to achieve a 5-star rating that reflect whether reporting measures in each domain has been achieved, particularly safety and mortality measures	Hospitals must report at least three measures in three groups; one group must be mortality or safety of care
7.	Balance the summary rating with the ability to drill down for more detailed information	
8.	Differentiate methodology from the user interface and enhance the user interface	
9.	Consider approaches to peer grouping that would be most meaningful to patients, for example, location, hospital characteristics, or service lines	Peer group by number of measure groups reported
10.	Expand the data sources to include a more comprehensive representation of patient populations served across hospitals	
11.	Consider aligning program design elements across programs when possible	Stratify the Readmission group by proportion of dual eligible patients (similar to HRRP)

The Hospital Quality Star Ratings are designed to summarize information from the existing measures on Hospital Compare in a way that is useful and easy to interpret for patients. Hospitals are assigned a

rating between one and five stars using publicly reported measure results. The ratings complement the HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) star ratings system.

The guiding principles of the Star Ratings are to utilize methods that are scientifically valid, inclusive of hospitals, account for heterogeneity of measures and hospitals reporting, and accommodate changes in underlying measures. CMS aims to maintain the transparency of the methodology and to continue to be responsive to stakeholder feedback. To further efforts to provide patients and caregivers access to quality information to inform decisions about their care, CMS also recently launched the Care Compare[1] site, which aggregates the eight CMS compare tools (including Hospital Compare) into one interface.

Proposed Methodology for 2021

High-level steps in the proposed 2021 Star Ratings methodology are as follows:

- Step 1: Select measures. Measure inclusion criteria are unchanged.
- Step 2: Group measures. Measures will be grouped into five groups, down from seven.
- Step 3: Calculate group score. Use a simple average of available measure scores rather than latent variable modeling.
- Step 4: Generate summary score. This step remains unchanged.
- Step 5: Apply reporting thresholds. Hospitals must report on three measures in at least three groups. One of these three groups must now be Mortality or Safety of Care.
- Step 6: Apply peer grouping. Classify hospitals into groups using number of measure groups reported with at least three measures (i.e., three, four, or five).
- Step 7: Calculating Star Ratings. K-means clustering within each peer group will be used to categorize summary scores into five-star ratings.

For this reconvening, the multistakeholder Panel reviewed the recently released proposed methodology for the Hospital Star Ratings[2] for calendar year 2021 to prepare for discussion on the following components: measure group construction, scoring approaches, and accounting for patient risk. Proposed changes were considered in the context of the goal of the Star Ratings program as decisions regarding methodology should reflect the overall intent of the Ratings and how they are being used. All Panel members were encouraged to submit written comments to CMS by October 5, 2020 as part of the rulemaking process.

Key Considerations from NQF Panel

This convening was a unique opportunity to bring diverse stakeholders together to have a joint dialogue and provide input on the evolution of the Hospital Star Ratings. After introduction and rationale for each proposed change was provided by CMS and Yale Center for Outcomes Research & Evaluation (Yale CORE) representatives, Panel members discussed whether they supported the proposed update and offered additional suggestions for CMS to consider in future methodology updates. Overall Panel members were generally accepting of a majority of the proposed updates and reacted positively to the responsiveness of CMS to stakeholder feedback. The Panel's dialogue around the proposed updates and considerations related to each topic are detailed in the following subsections.

Measure Selection and Grouping

Proposed reconstruction of measure groups

The current methodology groups measures into seven groups, including three process measure groups. CMS proposes to combine the three existing process groups (i.e., Effectiveness of Care, Timeliness of Care, and Efficient Use of Medical Imaging) into a new "Timely and Effective Care" group. Twelve

measures will be removed from the Star Ratings as a result of the Meaningful Measures Initiative[3]. The change would support a sufficient number of measures included in each group to allow for meaningful measure group scores. **Panel members were accepting of the proposed change to combine the three existing process measure groups but highlighted several points for CMS to consider and test when selecting and grouping measures in the future**.

1. Grouping measures based on clinical area. Several individuals concurred that the overall measure grouping approach is not reflective of the Panel's previous suggestion to group measures in a clinically-coherent way. For example, the new "Timely and Effective Care" group contains measures about patients with low-back pain, patients with metastatic cancer undergoing external beam therapy, and patients undergoing a colonoscopy. These measures apply to different patient populations and different clinicians provide the various procedures, and it is not clear that the combination of measures provides information of value to a patient making choices about a specific clinical condition or problem. To meet the goal of providing summary information to patients to inform decision making, it was suggested that grouping measures by clinical area is more patient focused and would align better with patients' needs when making care decisions. The U.S. News and World Report Hospital Rankings[4] were cited as an example of methodology that groups by clinical area. For example, a hospital's knee replacement score includes relative survival, prevention of prolonged hospitalizations, prevention of revision surgery, relative volume of patients, etc., serving as a comprehensive source of procedure-specific information for a prospective patient.
2. Retaining the influenza immunization measure. The measure incentivizes the importance of influenza vaccinations, especially among healthcare workers and during the COVID-19 pandemic.
3. Establishing a standard mechanism for measure removal. Continuing to streamline groups is key to achieving parsimony.
4. Including more measures in the overall ratings when feasible. Including more measures in the Hospital Star Ratings may better reflect the diversity of conditions treated and care provided at hospitals. It would also support moving towards a more clinically-cohesive strategy, which would require the inclusion of more condition-specific measures. If measure data collection and reporting becomes more automated, more measures could be included without imposing undue administrative burden.

Scoring Approaches

Proposed replacement of Latent Variable Model (LVM) with simple average methodology to calculate measure group scores

CMS proposes to use a simple average methodology, rather than latent variable modeling, to calculate measure group scores. The change would reduce the complexity of the methodology so it could be easily explained, and results could be used for quality improvement. For example, using a simple average method would help hospitals better understand which measure scores were driving their results and in which areas there is an opportunity to improve the group score by improving scores on specific measures. Since the proposed group scoring approach would no longer use a statistical model, the proposed rule indicates there would no longer be the need for score winsorization or to exclude measures with statistically significant negative loadings. The measure group scores would be standardized to account for different distributions of measure group scores that require rescaling.

The Panel supported the proposed change to use a simple average of measure scores rather than the LVM. They generally agreed it was responsive to stakeholder feedback, allows for greater understanding

of how scores are derived, and adds clarity to what the ratings assess. The change would allow hospitals to better understand the impact of their efforts towards improvement. It was suggested the change would also eliminate the dominance of one measure within a group. Several other points were raised and suggested for future review.

1. One Panel member suggested CMS reconsider eliminating measure score winsorization, citing some measures may still have some measurement error at an individual hospital level. Retaining measure score winsorization may be a way to account for such errors, thus avoiding penalizing hospitals unnecessarily.
2. Consider refining the weight of each measure. There may be better options than straight weighting to reflect the strategic prioritization of some measures over others. This approach could still allow transparency into how each measure is weighted, which is important to hospitals to determine focus areas for improvement and to patients in interpreting how to use the Star Ratings.
3. Carefully consider if simple is always better than complex as it relates to measurement. For example, complex methodology is often needed when assessing quality of care related to important outcomes like mortality and readmissions. Just because a methodology is not easily understood by consumers may not be a valid reason for not using it if it would be most appropriate.

Proposed updated reporting threshold for hospitals to receive a Star Rating

The proposed rule includes an update to require the reporting of three measures in three groups, one of which must be Mortality or Safety of Care to receive a Star Rating. This change is proposed to improve the face validity of the Overall Star Ratings and to ensure they reflect important aspects of quality of care for patients. It was highlighted in the proposed rule that based on January 2020 data, 125 hospitals did not report at least three measures in the Mortality or Safety of Care groups. These may be hospitals that previously would have received a Star Rating due to reporting, for example, a sufficient number of readmissions measures. The proposed update would mean these hospitals would no longer receive a Star Rating.

The Panel supported the proposed update to the reporting threshold to receive a Star Rating to ensure Star Ratings include measures in the priority areas of mortality or safety. The minimum number of measures and measure groups required to be reported has not changed but rather the change requires one group to be either Mortality or Safety of Care. The current methodology requires one group to be an "outcome" group (i.e., Readmission, Mortality, or Safety of Care). Several other suggestions were proposed for future consideration.

1. Consider a stricter reporting threshold. The Panel offered several suggestions: requiring measures from all groups to be reported, requiring hospitals to report on an adequate number of measures in the safety category, and setting a threshold (e.g., 60 percent) for the total number of measures that must be reported. These considerations would make the threshold to receive a rating more stringent but potentially add validity to the ratings by ensuring they consider quality broadly and include crucial outcome-focused topic areas. Several members supported requiring the patient safety measure group to be reported in order to receive a rating. One member emphasized that mortality measures are not a substitute for safety measures (as the proposed threshold update may imply). Others liked that safety measures are calculated using the entire eligible population and expressed safety is a cross-cutting, critical requirement of quality care. A Yale CORE representative advised that there are far fewer patient safety measures; therefore, to be more inclusive of hospitals that would receive a rating,

reporting on at least three measures in the patient safety or mortality groups is being proposed. The Panel questioned whether a distinctly better reliability score would be derived from a measure group with three measures reported instead of two, for example. A Yale CORE representative shared that previously, the decision was made based on both face validity and empirical testing. Signal to noise was calculated to determine how many measures were needed for appropriate variation considering the LVM approach. Testing supported that between two and 2.5 measures were needed, but face validity supported a threshold of three measures. One Panel member expressed that the decision should not justified only by face validity.

2. Expanding the data sources used for measure calculation. For example, the mortality measures are calculated using Medicare Fee-for-Service claims which represent a small fraction of the Medicare population at certain hospitals and in parts of the country. This suggestion aligns with the Panel's consideration from 2019 that including Medicare Advantage data would allow the Star Ratings to reflect quality results for a larger patient population.

Accounting for Patient Risk

Proposed stratification of the Readmission measure group only by hospitals' proportion of dual-eligible patients

The proposed rule considers stratifying the Readmission measure group based on the proportion of dual-eligible patients. "Dual-eligible" patients are those patients who are eligible for both Medicare and Medicaid[5]. This change is proposed in response to stakeholder concern (particularly provider concern) that hospitals face unique challenges preventing readmissions among patients with complex social risk factors. The proposed stratification approach aligns with the method used for the HRRP[6] payment determination. The readmission group scores would be stratified based on the national average group score and average group score within each dual-eligible group. Hospitals missing dual-eligible data would not be assigned to a dual-eligible group or receive an adjusted group score.

There was a positive response to stratifying the Readmissions measure group by proportion of dual-eligible patients since it is aligned with the method used in HRRP. While there is broad agreement regarding clinical risk adjustment of performance measures for public reporting when appropriate, whether social risk adjustment should be considered remains unresolved. Proponents of social risk adjustment argue that if it is not employed, programs may unfairly penalize providers who care for vulnerable populations and providers may be discouraged from providing care to vulnerable populations. Others urge caution about its use and the potential to create multiple standards of care. This also reflects the need for greater levers to address disparities. For the Hospital Star Ratings, the Panel expressed that stratification by proportion of dual-eligible patients seems to be an adequate proxy to use in the short term but social risk adjustment approaches must continue to evolve as the ability to capture social risk factor data improves.

There were several alternative perspectives shared or suggestions made by Panel members.

1. One member expressed concern that using dual-eligibility as a proxy for poverty or social risk may not accurately represent the hospital's population—a concern also related to the HRRP method. However, another member emphasized their support for using dual-eligible status as a proxy for risk, as persons who are dual-eligible are poorer financially, report worse health status, and have greater health care needs than other Medicare beneficiaries.
2. One member encouraged CMS to consider whether the risk stratification approach yields enough differences in hospital's ratings for it to be worth performing.
3. Consider how the stratification approach would have to be adjusted if measures are regrouped in the future. Adjustment at the individual measure level when conceptually and empirically

supported may be an alternative that would not warrant a new stratification approach if measure groups change over time.
4. Risk stratification may not be clear to consumers and may need to be carefully presented to support their understanding.

Proposed peer grouping methodology by number of measure groups reported

CMS is proposing to peer group hospitals based on the number of measure groups they report with at least three measures (i.e., three groups, four groups, or five groups). Hospitals would be grouped after their summary group scores are calculated and before they are assigned a Star Rating. This proposed methodology update aims to provide a more accurate comparison of quality between hospitals since hospitals would be compared to others within the same peer group. Hospitals report different numbers and types of measures that reflect differences in size, patient volume, case mix, and services provided. The peer grouping proposal is dependent on the inclusion of Critical Access Hospitals (CAHs), which make up the majority of the proposed "three measure" peer group. Historical data suggests most hospitals remain in the same peer group year over year (e.g., report a stable number of measure groups over time).

There was not general agreement supporting the proposed approach to peer group hospitals by number of measure groups reported. The Panel's main concern was that this approach is not intuitive and would not be easily understood by consumers, who are the target users of the Star Ratings. If this approach is finalized, the Panel encouraged CMS to carefully explain this peer grouping approach and how it should be interpreted with patients and consumers in mind. Some expressed there is not enough harmony in the characteristics of hospitals in each group; these groups may be more like clusters than peers with similar characteristics. While several members urged caution, at least one member was in support of this approach voicing that it is an elegant way to account for many different factors (including hospital size) and helps account for incomplete reporting and its impact on ratings. Since the number of measures reported is correlated well with hospital size, the proposed approach would allow comparisons between similar sized hospitals and reduce the impact of "crowding out" of hospitals based on size—an update they consider would provide better, more transparent information to consumers.

Several considerations were raised for CMS to consider and test in the future.

1. Examine how these peer groups overlap with clinical service lines and other hospital characteristics.
2. Examine alternative approaches that peer group hospitals and compare performance based on clinical areas or services provided. This approach was also suggested during the Panel's previous convening in 2019 and would likely be of interest to patients who search for care based on their specific condition or the service they are seeking (and geographic region).

Next Steps

CMS will consider comments submitted as part of their rulemaking process and finalize the Hospital Star Ratings methodology for the 2021 calendar year. The NQF Panel expressed agreement with most of the proposed updates and offered several alternative considerations for future exploration. Many proposed changes to the Star Ratings methodology directly address stakeholder feedback and a number of them align with considerations from NQF's 2019 Summit, an encouraging example of collaboration within the quality measurement industry.

This work demonstrates how multistakeholder recommendations can advance the design of quality reporting and payment programs. Measurement systems are essential for assessing and paying for

quality and value, supporting consumer decision-making, and driving improvement. While methods to assess the credibility of individual performance measures are well established and the value of their endorsement is widely accepted, there is no process in place for reviewing the measurement systems in which they are used. NQF's Star Ratings convenings are an early example of such a review. Furthermore, recent work by NQF's Measure Sets and Measurement Systems Technical Expert Panel puts forth a preliminary framework for their evaluation, which may help fill an existing gap in the measurement infrastructure and healthcare quality ecosystem.

References

1 Centers for Medicare and Medicaid Services. Find Healthcare Providers: Compare Care Near You. Medicare.gov. https://www.medicare.gov/care-compare/. Last accessed October 2020.

2 Centers for Medicare and Medicaid Services. *85 Federal Register 48772, Page 48772-49082 (Proposed August 12, 2020)*. https://s3.amazonaws.com/public-inspection.federalregister.gov/2020-17086.pdf.

3 Centers for Medicare and Medicaid Services. Meaningful Measures Hub. CMS.gov. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/MMF/General-info-Sub-Page. Published September 10, 2019. Last accessed October 2020.

4 U.S. News & World Report. U.S. News & World Report Hospital Rankings & Ratings. https://health.usnews.com/best-hospitals. Last accessed October 2020.

5 Centers for Medicare and Medicaid Services. Medicare Learning Network: Dually Eligible Beneficiaries Under Medicare and Medicaid. February 2020. https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/medicare_beneficiaries_dual_eligibles_at_a_glance.pdf.

6 Centers for Medicare and Medicaid Services. Hospital Readmissions Reduction Program (HRRP). CMS.gov. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program. Last accessed October 2020.

Appendix A. Participants

Expert Panel Member	Title	Organization
Nancy Foster, AB	Vice President for Quality and Patient Safety Policy	American Hospital Association
Cristie Upshaw Travis, MSHHA	Chief Executive Officer	Memphis Business Group on Health
Lisa Freeman, BA	Executive Director	Connecticut Center for Patient Safety
Lindsey Galli, BA	Patient Family Advisor and Director of Education	Patient and Family Centered Care Partners
Danielle Lloyd, MPH	Senior Vice President, Private Market Innovations and Quality Initiatives	American's Health Insurance Plans
David Nerenz, PhD	Director Emeritus of the Center for Health Policy and Health Services Research	Henry Ford Health System
Jody Amodeo, BSN	Senior Consultant	Willis Tower Watson
Andy Amster, MSPH	Director of the Center for Healthcare Analytics	Kaiser Permanente
John Matthew Austin, PhD, MS	Assistant Professor of Anesthesiology and Critical Care Medicine	Johns Hopkins University, Armstrong Institute for Patient Safety and Quality
Leah Binder, MA, MGA	President and CEO	Leapfrog
John Bott, MBA, MSSW	Consultant	Self-employed
Stephanie Clouser, MS	Data Scientist	Kentuckiana Health Collaborative
Maryellen Guinan, JD	Senior Policy Analyst	America's Essential Hospitals
Ben Harder, BA	Managing Editor and Chief of Health Analysis	U.S. News and World Report
Bala Hota, MD	Vice President, Chief Analytics Officer	Rush University
Omar Lateef, DO	President and CEO	Rush University Medical Center
Wendy Marinkovich, BSN, MPH, RN	Executive Director, Provider Measurement Programs	Blue Cross Blue Shield
Kristine Martin Anderson, MBA	Executive Vice President	Booz Allen Hamilton
Janis Orlowski, MD, MACP	Chief Healthcare Officer	Association of American Medical College
Louise Probst, BSN, MBA	Executive Director	St. Louis Area Business Health Coalition
Karen Roth, RN, MBA, CPA	Director of Research	St. Louis Area Business Health Coalition
Clarke Ross, DPA, MA	Public Policy Director	American Association on Health and Disability
Claudia Salzberg, PhD, MSE	Vice President, Quality	Federation of American Hospitals
David Shahian, MD	Vice President for Quality	Massachusetts General Hospital

CMS Representatives	Title
Joseph Clift, EdD, MPH, MS, PMP	Acting Manager, Division of Quality Measurement
Michelle Schreiber, MD	Group Director of Quality Measurement and Value-Based Incentives Group
Reena Duseja, MD, MS	Chief Medical Officer for Quality Measurement
Annese Abdullah-Mclaughlin, BSN	Public Reporting and Star Ratings Lead

Yale CORE Representative	Title
Arjun Venkatesh, MD, MBA, MHS	Research Scientist

National Quality Forum Staff	Title
Shantanu Agrawal, MD, MPhil	President and CEO
Sheri Winsper, RN, MSN, MSHA	Senior Vice President, Quality Measurement
Maha Taylor, MHA, PMP	Managing Director
Nicolette Mehas, PharmD	Director
Teresa Brown, MHA, MA, CPHQ, CPPS	Senior Manager
Yvonne Kalumo-Banda, MS	Manager