Summary of Technical Expert Panel (TEP): Overall Hospital Quality Star Ratings on Care Compare

October 2023

Table of Contents

Background

The Centers for Medicare & Medicaid Services (CMS) contracted with Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE) to maintain the Overall Hospital Quality Star Rating on Care Compare (hereafter referred to as "Overall Star Rating"). The contract name is Development, Reevaluation, and Implementation of Outcome/Efficiency Measures for Hospital and Eligible Clinicians, Option Period 4. The contract number is HHSM-75FCMC18D0042, Task Order Number HHSM-75FCMC19F0001. As part of its measure development and maintenance process, CORE convenes groups of stakeholders who contribute direction and thoughtful input to the measure developer during project refinement and maintenance.

The primary goal of this Technical Expert Panel (TEP) is to support the maintenance and evolution of the Overall Hospital Quality Star Rating methodology. This approach is consistent with CMS's approach for iterative improvement of quality measures and quality programs. The Overall Hospital Quality Star Rating project is designed to create a summary of the hospital quality measures publicly reported on CMS's Care Compare website so that these aspects of quality are presented in a meaningful and accessible way to patients and consumers.

CMS and the CORE project team have sought to be transparent and responsive to stakeholder input throughout development and reevaluation of the Overall Star Rating. During the initial development of the methodology, CORE convened a multi-stakeholder TEP and Person & Family Engagement Workgroups (PFE WG), held two public input periods, hosted two National Provider Calls, and performed a hospital dry run before launching the Overall Star Rating in 2016. Since 2016, CMS and the development team have hosted four National Provider Calls, held nine PFE Workgroup meetings and nine Provider Leadership Workgroup (Provider WG) meetings. This 2023 Call for TEP marks the fourth iteration of the Overall Hospital Quality Star Ratings TEP membership and includes some new members in addition to some continuing members; previous iterations of the TEP met eight times since 2014.

The TEP currently includes 24 experts in consumer perspectives, purchaser perspectives, quality improvement, performance measurement, statistical modeling and empirical methods, healthcare disparities, and patient/family/caregiver perspectives, who provide input on key methodological decisions during the reevaluation of the Overall Star Rating.

This report summarizes the feedback and recommendations provided by the current iteration of the TEP during its first meeting on October 31st, 2023.

Measure Team

The CORE Overall Star Rating development and reevaluation project team is led by Dr. Cameron Gettel and Mr. Kyle Bagshaw and overseen by Project Director Dr. Arjun Venkatesh. See Table 1 for the full list of CORE team members on the Overall Star Rating team.

Table 1. Star Ratings Team Member List

Name	Role
Arjun Venkatesh, MD, MBA, MHS	Project Director Chair and Chief, Department of Emergency Medicine at Yale School of Medicine and Yale New Haven Hospital; Scientist, Yale CORE
Kyle Bagshaw, MPH	Project Co-Lead
Cameron Gettel, MD, MHS	Project Co-Lead
Li Qin, PhD	Lead Analyst
Shefali Grant, MPH	Project Manager
Eve Rothenberg, BA	Project Coordinator
Prince Omotosho, BS	Research Support
Jeph Herrin, PhD	Statistical Consultant
Zhenqiu Lin, PhD	Data Management and Analytics Director
Lisa Suter, MD	Senior Director
Roisin Healy, BA	Person and Family Engagement Team Coordinator
Mariel Thottam, MS, BCBA	Person and Family Engagement Team Lead
Thushara John, MA, MHA	Person and Family Engagement Team Lead
Ariel Williams, MPH	Person and Family Engagement Research Support

Technical Expert Panel

In alignment with the CMS Measures Management System, and under the guidance of CMS, CORE held a 30-day public call for nominations and convened a TEP for the reevaluation of the Overall Star Rating. The role of the TEP is to advise CORE on key methodological and analytical updates; CORE then summarizes and considers this input when making recommendations to CMS for its final reevaluation decisions on the Overall Star Rating. Convening the TEP is one important step in the process that ensures transparency and provides an opportunity to obtain balanced input from multiple stakeholders.

CORE solicited potential TEP members via email by contacting individuals and organizations represented on the previous TEPs or recommended by the measure team; email blasts sent to the CMS physician and hospital email listservs; and through a posting on CMS's website. After reviewing TEP nominations, CORE

confirmed a TEP of 24 members, listed below. The current TEP appointment term is from October 2023 to March 2024, with the possibility of extending to additional meetings and future years.

Table 2. Star Ratings TEP Member List

Participant and Credentials	Title	Organization, State
Amy Chin, MS	Hospital Administrator	Hospital for Special Surgery, NY
Amy Minnich, RN, MHSA, CPC	Associate Vice President, Quality Safety and Patient Experience	Geisinger, PA
Ashantae Okechukwu	--	IL
Benjamin D. Pollock, PhD, MSPH	Assistant Professor of Health Services Research, Scientific Director for Data Science	Mayo Clinic, FL
Carol B. Pugh, PharmD, MS	Retired Clinical Pharmacist/Biostatistician	VA
David Levine, MD, FACEP	Group Senior Vice President, Advanced Analytics & Data Science	Vizient, Chicago, IL
David M. Shahian, MD	Vice President, Center for Quality & Safety	Mass. General Hospital, Society of Thoracic Surgeons, MA
Danielle Rosario	HR Business Partner	AZ
Erin O'Rourke, BS	Executive Director, Clinical Performance and Transformation	America's Health Insurance Plans (AHIP), DC
Eugene Hsu, MD, MBA	Physician Executive in Clinical Quality	Elevance Health, CA
Itisha S. Jefferson, BA	Patient/Advocate/Caregiver	IL
Jennifer Lamprecht, MS, RN, CNL, CPHQ	Director of Quality Strategy	Sanford Health, SD
John Bott, MSSW, MBA	Independent Contractor	N/A
John Martin, PhD, MPH	Vice President, Data Science	Premier, Inc., NC
Jordan Russell, MPA, CPHQ	Executive Director of Quality	UnityPoint Health, IA
Julie Wall, RN, MBA, FACMPE	Senior Vice President, Quality & Patient Safety	Benefis Health System, MT

Participant and Credentials	Title	Organization, State
Kacie Kleja, MBA, MS	Vice President, Clinical Data & Analytics	HCA Healthcare, TN
Lisa Freeman, BA	Patient, Caregiver, Executive Director	CT Center for Patient Safety, CT
Karl Bilimoria, MD, MS	Chair, Department of Surgery, System Vice President for Quality, Executive Director of the Surgical Outcomes and Quality Improvement Center (SOQIC)	Indiana University School of Medicine, IN
Larry Boress, MPA	Executive Director	National Association of Worksite Health Centers, AZ
Marisha Burden, MD, MBA	Division Head of Hospital Medicine, Professor of Medicine	University of Colorado School of Medicine, CO
Roger Lacoy	Patient, Caregiver, Advocate, Board of Directors	PFCC Partners, PITCH and PHC Inc., IA
Sandi Hyde, BSME, MSPS	AVP Clinical and Operations Quality Data and Regulatory Reporting	Lifepoint, TN
Stephanie Fishkin, PhD	Data Analytics and Reporting Consultant	Kaiser Permanente, CA

Specific Responsibilities of the TEP members

Specific responsibilities of TEP members include:

- Complete and submit all nomination materials, including the TEP Nomination Form, letter of interest, disclosure of conflicts of interests, and curriculum vitae;
- Review background materials provided by CORE prior to each TEP meeting;
- Attend and actively participate in the TEP in-person meeting and/or teleconference meeting(s);
- Provide input and feedback to CORE on key clinical, methodological, and other decisions;
- Provide feedback to CORE on key policy or other non-technical issues;
- Review the TEP summary report prior to public release; and
- Be available to discuss recommendations and perspectives following group TEP meetings and public release of the TEP summary report.

CORE provides an agenda and background materials before every meeting for TEP members to review. TEP members are generally expected to attend a majority of meetings, and to review and comment on materials for the meetings they cannot attend. CORE then summarizes member comments and recommendations in a report that will be publicly posted on CMS's website.

TEP Meeting

The original TEP participated in three meetings from 2014 through 2015. The second iteration of the TEP participated in four meetings from 2017 through 2018, and the third iteration of the TEP participated in one meeting in 2019. The discussion at the fourth iteration of the TEP presented in this report was held on October 31st, 2023.

Throughout this report, references to the TEP or to TEP meetings refer specifically to the current iteration of the TEP, unless specified otherwise.

TEP meetings follow a structured format consisting of the presentation of key issues identified during measure development or reevaluation, as well as CORE's potential approaches to addressing the issues, followed by an open discussion of these issues by the TEP members.

TEP Overview

Prior to the first TEP meeting, CORE provided TEP members with a contextual document describing general project background information as well the detailed steps of the Overall Star Rating methodology. The TEP was also provided with the PowerPoint presentation that would be presented live at the meeting.

The goals of this TEP meeting were to present, discuss, and gain feedback from the TEP on potential updates to the Overall Star Rating methodology that would aim to emphasize the Safety of Care measure group.

The following bullets represent a **high-level summary** of what was discussed during the first TEP meeting.

Background and Approach

- CORE reviewed the Star Ratings background and methodology, noting the objective was to summarize publicly reported CMS hospital quality measures in five Measure groups (Mortality, Safety of Care, Readmission, Patient Experience, and Timely & Effective Care) into one overall hospital rating for reporting on CMS's Care Compare website.
- CORE presented information about the relationship between the Safety of Care measure group and Overall Star Rating performance, and noted CMS is increasing its efforts around Safety of Care across all programs by emphasizing it in the National Quality Strategy and Universal Foundation Measures.
- CORE provided background on Safety of Care in the Star Rating, noting there are currently eight measures in the Safety of Care measure group; adding, removing, or altering the Safety of Care measures is outside of CORE's scope, but CORE is able to implement methodology updates at the measure group level or Overall Star Rating level.
- CORE noted there are a small number of hospitals with poor Safety of Care performance that still achieved a 5-Star Rating in the July 2023 Star Rating release and presented about potential options to minimize or eliminate the number of hospitals that performed poorly on Safety of Care but still achieved 5 Stars. These options include reweighting the Safety of Care measure group to 30% and/or application of a Star Rating penalty/cap to prevent hospitals in the lowest quartile of performance for Safety of Care from achieving 5 Stars.
- CORE noted the importance of the stakeholder input provided by the PFE WG, Provider WG, and the TEP in the evolution of the Star Ratings since their inception in 2015. CORE shared specific feedback provided earlier in October 2023 by the PFE WG, and requested the TEP's input on the options of reweighting Safety of Care and applying a Star Rating penalty/cap.

Summary of TEP Input (including both Zoom and written responses)

- Overall, the TEP feedback suggested varying support for the options proposed (reweighting or a Star Rating cap/penalty) and more exploration of the alternative ways to elevate Safety of Care.
    - Participants broadly agreed on the importance of Safety of Care as a concept and that it might be elevated using a different approach.
    - Several participants noted the importance of a consistent and fair strategy for scoring hospitals that minimizes hospitals being penalized multiple times for poor Safety performance.
    - Some participants noted concerns with alignment of this program with other programs such as the Hospital-Acquired Condition (HAC) Reduction Program and Hospital Value-Based Purchasing (HVBP).
    - Several participants expressed concern about how such changes to the Star Rating methodology might be communicated on Care Compare.

Reweighting

- Specific TEP feedback regarding changing the relative weights of the Measure Groups (Mortality, Safety of Care, Readmission, Patient Experience, and Timely & Effective Care) in the Star Rating are included the below:
    - A few TEP members noted that reweighting seems like a reasonable way to emphasize Safety within the current methodology.
    - Many TEP participants expressed concerns about reducing the weights of the other measure groups, and particularly about giving a greater emphasis to Safety of Care than Mortality.
    - Several TEP members expressed concerns that hospital performance across the measure groups was interrelated, with performance in Safety of Care also impacting the other measure groups, such as Readmission or Patient Experience; and conversely, the other measure groups such as Readmission also conveying information about Safety of Care.
    - Some TEP members conveyed concerns about how the proposed changes would impact smaller hospitals that do not have enough predicted events to calculate a Standard Infection Ratio (SIR) for the HAC measures, whose measure results could look very different than larger hospitals; other participants noted concerns related to hospitals with different characteristics, for example that academic teaching hospitals which treat sicker patients may already have lower Star Ratings.
- Some of the participants opposed the reweighting option as it was currently proposed, with additional feedback suggesting:
    - Many participants noted that the proposed measure group weights seemed arbitrary.
    - Several participants voiced a preference for more evenly weighting the measure groups.
    - Some participants expressed uncertainty that reweighting would drive changes in hospital behavior/performance.

Star Penalty/Cap

- TEP feedback supporting a Star Rating penalty/cap included suggestions that hospitals performing poorly in any measure group should not be able to achieve a 5-Star Rating, with multiple participants recommending low scores in any measure group should be considered if a penalty/cap is adopted. However, TEP members communicated several concerns about application of a Star Rating penalty/cap:

- Some TEP members thought a cap/penalty may penalize hospitals with poor performance twice, as they may already have a low Star Rating in part because of that performance; furthermore, these hospitals may be additionally penalized through other CMS programs (such as HAC Reduction).
- Some TEP members expressed that applying a cap based solely on performance in a single measure group, or when a hospital reports fewer than three measures (or fewer than half of the measures) in each of the measure groups, could risk unduly penalizing hospitals based on too little information.
- Some TEP members questioned using the lowest quartile of performance, as opposed to the lowest quintile or decile; some recommended that a test of significance be applied to ensure performance is significantly lower than for other hospitals before applying a cap/penalty.

Comparing Presented Options & Additional TEP Suggestions

- Several TEP members expressed a preference for a policy-based penalty/cap option over reweighting preferences for reweighting vs. implementing a Star Rating penalty/cap included the below:
  - Of those who expressed a preference, there was a slight leaning toward a cap/penalty over reweighting, with several leaning toward a cap over a penalty.
  - However, a few participants expressed a preference for a penalty as it would impact the Star Ratings of a larger group of hospitals, and there was even some support for imposing a penalty greater than one star.
- A few TEP comments supported reweighting over a cap/penalty option because reweighting would reflect performance across all hospitals, rewarding those with excellent Safety of Care performance rather than only penalizing those with low Safety of Care performance.
- The TEP offered additional comments and suggestions, including:
  - Some participants felt all the options presented seem "manipulative" of the whole system and that CMS should similarly consider all measure groups rather than singling out Safety of Care (particularly since the discussion appears motivated by a very small subset of hospitals with good Star Ratings despite poor Safety).
  - Several participants expressed concerns about the current set of measures in the Safety of Care group, with several participants noting they fall short of representing safety, as they do not include a large enough population of patients and may not fully reflect the true safety of hospital. In particular due to low volume, many hospitals can only report on a subset of the measures and their scores might be disproportionally affected by a single event.
    - Some of these participants suggested adding additional measures to the Safety of Care measure group; the measure group might be more representative of services provided to make this information more relatable for consumers.
    - One participant recommended that if measures were going to be added, it would be ideal to begin with a theoretical framework of what should be in it to capture Safety holistically, rather than simply adding on to the existing measures.
    - Another participant suggested including a measure(s) of psychological and/or emotional safety, in addition to the current measures for physical safety.

- Several TEP members noted concerns about patients' interpretation of a 4-Star or 5-Star rating.
    - A few TEP members felt patients should not have to decide between a good surgical outcome or a safe place because every hospital should be safe, and the quality of outcomes is a separate issue.
    - Patients would likely assume 5-Star hospitals to have high scores in all the measure groups; some TEP members felt hospitals with poor Safety of Care performance should not have 5 stars and further that hospitals in the lowest decile for any measure group should not be awarded 5 Stars.
        - Similarly, patients may assume that even a 4-Star hospital would not have low scores for Safety of Care or Mortality.
    - TEP participants noted the importance of clear messaging on Care Compare so patients know a cap/penalty was applied and suggested that more information could be included such as the peer grouping that reflects hospital case volume.
        - Alternatively, one participant recommended addressing this concern by allowing consumers to see measure group contribution to the Overall Star Rating of a hospital, such as a visual representation where the consumer could readily see that a 5-star hospital was "green" in most measure groups, but "red' in Safety as this might be more meaningful.
    - Several participants noted that there is a small number of "outlier" hospitals with poor Safety but 5-star ratings and were concerned about making broad methodologic changes to address this small group of hospitals.
- A few participants noted broader methodological concerns about the current Star Ratings methodology, including:
    - The use of z-scores to standardize measure group scores that include measures with skewed distributions.
    - Questions about the stability of the Star Rating over time for a given hospital, and hesitance to support methodology changes based on observations for a single performance period.
- Several comments recommended additional stakeholder engagement, including communication with a subset of the lower performing hospitals to ensure there is a clear diagnosis of the problem before making methodology changes; engagement with a patient focus group to better understand how patients perceive Safety of Care in Star Ratings; and engagement of an expert panel to decide on weighting of the various measure groups rather than the presented reweighting to support a CMS priority at a given point in time.
- Additional suggestions included the addition of a workforce measure, such as stability or staffing, and further exploration of the relationships between price transparency and variation with the Overall Star Rating.

Next Steps

Ongoing Reevaluation

The project team will consolidate the feedback received at the October 31st, 2023, TEP meeting with the feedback received at the Person & Family Engagement and Provider Leadership Workgroup meetings in

October and November, respectively. The project team will share all stakeholder engagement feedback for CMS to consider in their decision-making regarding Safety of Care reevaluation in the Overall Star Rating.

Conclusion

The TEP provided important feedback for CORE to consider for potentially emphasizing Safety of Care in the Overall Star Rating. CORE presented three options for adjusting the methodology (reweighting, Star cap and Star penalty) and the TEP provided feedback on all three options, along with additional considerations related to the Safety of Care group specifically, to the other measure groups, and to the Overall Star Rating more broadly. While the TEP did not establish a clear consensus in support of any proposed option, the meeting produced a robust discussion of potential advantages and disadvantages of each along with other considerations; CORE will closely consider all feedback in subsequent analyses and recommendations to help inform CMS's ultimate decisions.

Appendix A. TEP Call Schedule

TEP Meeting #1

Tuesday, October 31, 2023 – 2:30-4:30 EST (Zoom teleconference)

Additional TEP meetings to be determined based on Yale CORE need and TEP availability.

Appendix B. Detailed Summary of HOP4 TEP Meeting #1

Tuesday, October 31st, 2023, 2:30 – 4:30 PM EDT

Participants

Technical Expert Panel (TEP) Members:

Karl Bilimoria, MD, MS; Larry Boress, MPA; John Bott, MBA, MSSW; Marisha Burden, MD, MBA; Amy Chin, MS; Stephanie Fishkin, PhD; Eugene Hsu, MD, MBA; Sandi Hyde, BSME, MSPS; Itisha S. Jefferson, BS; Kacie Kleja, MBA, MS; Roger A. Lacoy; Jennifer Lamprecht, MS, RN, CNL, CPHQ; John Martin, PhD, MPH; Amy Minnich, RN, MHSA, CPC; Erin O'Rourke, BS; Benjamin D. Pollock, PhD, MSPH; Carol B. Pugh, PharmD, MS; Danielle Rosario; Jordan Russell, MPA, CPHQ; David M. Shahian, MD.

Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE):

Kyle Bagshaw, MPH; Patricia Faraone Nogelo, MSW, PhD, LCSW; Cameron Gettel, MD, MHS; Shefali Grant, MPH; Roisin Healy, BA; Prince Omotosho, BS; Eve Rothenberg, BA; Mariel Thottam, MS, BCBA; Arjun Venkatesh, MD, MBA, MHS.

X4 Health:

Stephanie Lambert, BA.

Veterans Health Administration (VHA) [Non-Voting Guests]:

Rachael Hasselbeck, RN; Jim Krabacher, BS.

Detailed Discussion Summary

Welcome, Introductions, and Meeting Objectives

- Ms. Mariel Thottam introduced herself as the Stakeholder Engagement Team Lead for the Yale New Haven Health Services Corporation–Center for Outcomes Research and Evaluation (CORE), welcomed the group and reminded members that materials and discussions of the Technical Expert Panel (TEP) are confidential and should not be disseminated or discussed outside of the group until made public by the Centers for Medicare and Medicaid Services (CMS). She stated the TEP is funded by CMS through a contract with CORE.
- Ms. Thottam provided an overview of CORE's work on outcome measure development, noting the importance of engaging with patients and stakeholders directly to better inform their measure development work.
- Ms. Thottam noted the Star Ratings TEP was originally convened in 2015, and the current TEP included some original members along with some new members. She noted the TEP currently includes 24 members representing a variety of perspectives (e.g., patients, caregivers, advocates, clinicians, health plans, and directors of hospitals/health systems) from across the nation. She

reminded TEP members they represent themselves on the TEP, rather than their organizations, and she requested TEP members keep the CORE team apprised of any changes in their affiliations and conflicts of interest.

- Ms. Thottam conducted a roll call of the TEP participants with 20 of the 24 members being present.
- Ms. Thottam reviewed the meeting objectives including:
    - Review Star Ratings Background and Methodology;
    - Presentation of the Background on Safety of Care in the Star Rating;
    - Presentation of Reevaluating Safety in the Star Rating;
    - Discussion of Options for Reevaluating Safety in the Star Rating;
    - Open Discussion (time permitting); and
    - Next Steps and Meeting Conclusion.

Review Star Ratings Background and Methodology

- Dr. Gettel thanked the TEP members for their interest in this work and reviewed the Star Rating project overview and background, including the guiding principles.
    - The project objective is to summarize Medicare quality information reported on the Care Compare website in a single summary Star Rating to optimize use for patients.
    - The project timeline spans from 2015 to present and has included a variety of stakeholder engagement activities including Patient & Family Engagement Workgroups (PFE WG), Provider Leadership Workgroups, and Technical Expert Panels (TEP).
    - The Overall Star Rating is updated annually with the most recent release having occurred in July 2023, based on January 2023 data. This was the first Star Rating release in which Veterans Health Administration (VHA) hospitals were eligible to receive a Star Rating.
    - There are seven steps in the Overall Star Rating methodology:
        1. Selection and standardization of publicly reported measures;
            - To be included a measure must be required by a CMS program and publicly reported on Care Compare
        2. Assignment of measures into measure groups (Mortality, Safety of Care, Readmission, Patient Experience, and Timely & Effective Care);
        3. Calculation of measure group scores as simple average of measure scores;
            - Measure group scores are then themselves standardized.
        4. Generation of summary score as weighted average of measure group scores;
            - Each group has a base weight of 22%, except for Timely & Effective Care at 12%;
            - In the event that a hospital has measures in one of the five measure groups, their score is redistributed proportionally to the remaining groups.
        5. Application of reporting thresholds;
            - Hospitals must have reported at least three measures in at least three measure groups, one of which must be either the Mortality or Safety of Care, to receive a Star Rating.
        6. Peer grouping is applied based on the number of measure groups with at least 3 measures included in the calculation of the Star Rating;
            - For example, a hospital that had three or more measures reported in all five measure groups is included in the '5 Measure Group Peer Group', while a

hospital that had three or more measures in just three measure groups is included in the '3 Measure Group Peer Group'. Most rated hospitals fall within the '5 Measure Group Peer Group.'

7. Within each peer group, hospitals are assigned a rating between 1 and 5 stars using a k-means clustering algorithm.

- Dr. Gettel noted that adding, removing, or altering Safety of Care measures is beyond the established measure inclusion and exclusion criteria and is outside of CORE's scope; however, CORE can implement methodology updates at the measure group level or Overall Star Rating level if aligned with the original principles of the program.

Presentation: Background on Safety of Care in Star Ratings

- Mr. Bagshaw shared additional information about the relationship between Safety of Care and Star Rating performance, noting CMS is increasing efforts around Safety of Care across all programs by emphasizing it in the National Quality Strategy (NQS) and Universal Foundation of Measures. Safety data is reported for Hospital Associated Infections (HAI) through the Centers for Disease Control and Prevention National Healthcare Safety Network, with scores used in the CMS Hospital-Acquired Conditions (HAC) Reduction Program.
- Mr. Bagshaw noted the Safety of Care measure group has been included in the Star Rating since its inception, always weighted at 22% of the total Star Rating score.
    - Beginning in 2021, Safety of Care and Mortality were further emphasized in the Star Rating methodology by requiring hospitals to receive at least three publicly reported measure scores for at least three measure groups, one of which must be Mortality or Safety of Care. This decision was partially informed by stakeholder feedback emphasizing the relative importance of Mortality and Safety of Care.
    - There are eight measures in the Safety of Care measure group, including six HAI measures (HAI-1 – HAI-6), one Complications measure after total hip or total knee replacement (Hip/Knee), and one adverse event measure composite measure (PSI-90).
    - Reporting status of each measure varies depending on whether the hospital meets the minimum case counts threshold for the measure; in general hospitals are required to collect and report data elements for all Star Rating measures and do not have the ability to select which of their measure scores are reported.
- Mr. Bagshaw reviewed the percentage of hospitals reporting each of the eight Safety of Care measures by peer group designation.
    - There were 194 Peer Group 3 hospitals, 462 Peer Group 4 hospitals, and 2,420 Peer Group 5 hospitals represented.
    - Seven of eight measures are reported by a large majority of Peer Group 5 hospitals based on sufficient volume, while reporting is much lower among Peer Group 3 and 4 hospitals.
- Dr. Gettel noted there was a strong relationship between the Safety of Care measure group and the Overall Star Rating, with hospitals that did well in Safety of Care tending to also do well on the Star Rating, but there were a few outliers that had lower Safety of Care outcomes that achieved a 5-Star Rating.
    - These outliers do not align with CMS's emphasis on Safety of Care.
    - Poor performance was defined as being in the bottom quartile of performance, or the lowest 25% of measure group scores.

- Hospitals with bottom-quartile Safety of Care scores have predominantly below-average performance in all individual Safety of Care measures.
- In July 2023, there were 748 hospitals in the bottom quartile for the Safety of Care score and 20 of those hospitals still achieved a 5-star rating.
- CORE's goal is to decide how to manage these situations with poor Safety of Care outcomes and 5-star ratings.

- Dr. Gettel shared a graph showing the cumulative star rating measure group scores for the 20 hospitals with poor performance on Safety that achieved 5-star ratings. These hospitals tended to perform very well in the other measure groups (Mortality, Readmission, Patient Experience, and Timely & Effective Care).

Presentation: Reevaluating Star Ratings

- Dr. Gettel noted two potential options to achieve CORE's goal of minimizing or eliminating the number of hospitals that performed poorly on Safety of Care but still did well in the Star Rating.
  - Option 1 was reweighting, or increasing the weight of Safety of Care and proportionally reducing the other measure group percentages to ensure the ratios between the other domains would stay the same, to highlight the importance of Safety of Care.
  - Option 2 was applying a Star Ratings penalty or cap, by applying a policy that explicitly limits how well a hospital could perform in Star Rating if they are in the bottom quartile of the Safety of Care measure group.

Option #1 — Reweighting

- Dr. Gettel discussed option 1, noting CORE evaluated the 2023 Star Ratings data using reweighting. A comparison of current weights and modified weights under reweighting are included in the table below:

Measure Score Group	Current Weight	Reweighting
Safety of Care	22%	30%
Mortality	22%	19.7%
Readmissions	22%	19.7%
Patient Experience	22%	19.7%
Timely & Effective Care	12%	10.8%

  - Reweighting the July 2023 Star Ratings resulted in three hospitals with low Safety of Care results achieving a 5-Star Rating, as compared with 20 hospitals with 5-star ratings using the original/current methodology. Overall, 233 of 3076 hospitals would receive a lower Star Rating and 213 would receive a higher rating with the proposed reweighting compared to the current weight.
  - Pros of reweighting include:
    - Consistent with the foundation of the current Star Ratings methodology;
    - Would not require additional external data;
    - Reduces the number of hospitals with poor performance on Safety of Care achieving a 5-star rating; and
    - Emphasizes good performance in Safety of Care.
  - Cons of reweighting include:

        - Does not eliminate 5-star rating hospitals with poor Safety of Care performance; and
        - Slightly reduces the influence of the other measure groups.

Option #2 — Star Penalty/Cap Background

- Mr. Bagshaw discussed option 2, noting CORE is considering applying a policy-based penalty or cap to the Star Rating that would limit how well a hospital could perform overall if they have poor Safety of Care outcomes.
    - A Star Penalty (Option 2a) that would reduce the star rating of any hospital in the worst quartile of Safety of Care by one star; or
    - A Star Cap (Option 2b) that would enforce a limit for hospitals in the worst quartile on Safety of Care to a maximum of 4 Stars. In this approach, if a facility has a 2, 3, or 4-star rating, their score would not be reduced, but if they received a 5-star rating, then they would be reduced to a 4-star rating.
    - Either the Star Penalty or Star Cap could apply to any hospitals that reported at least one safety measure, or only those that reported at least three safety measures. A group score based off just 1 or 2 measures may be subject to greater variation and may not be consistent with goal of the Star Rating to capture a snapshot of overall performance. For the analyses presented, the penalty/cap was applied only for hospitals with at least three safety measures.
    - Pros of a Star Penalty include:
        - It emphasizes Safety of Care through a new standard for all hospitals, regardless of their Star Rating.
    - Cons of a Star Penalty include:
        - Most hospitals with poor Safety of Care already had a poor Star Rating;
        - Hospitals with poor Safety of Care that already receive a 1-Star Rating are effectively exempt;
        - Some hospitals that performed excellently in all other measure groups would still be penalized; and
        - Hospitals could be penalized based on poor performance on only a single Safety of Care measure.
    - Pros of a Star Cap include:
        - It is a more targeted solution to the issue of hospitals performing poorly in Safety of Care receiving 5 Stars; and
        - The cap applies equally to all hospitals: the 5-star rating is reserved for hospitals that achieved a minimum threshold in Safety of Care.
    - Cons of a Star Cap include:
        - Some hospitals that performed excellently in all other measure groups would still be penalized; and
        - Hospitals could be penalized based on poor performance of only a single Safety of Care measure.

Summary of Options

- Mr. Bagshaw reviewed both the Reweighting option and the Penalty/Cap options. He noted both options could be considered independently, and reweighting is not mutually exclusive with the

penalty/cap approach. The penalty is more explicit and stronger in terms of the emphasis on Safety of Care than reweighting, preventing hospitals with poor Safety of Care from achieving a high star rating.

- Ms. Thottam asked the TEP participants if they had questions about the material the CORE team presented.
- A TEP member asked whether there were other measure groups where hospitals perform in the worst performing quartile and receive a 5-star rating.
    - Dr. Gettel responded that CORE evaluated this for the Mortality measure group and observed similar numbers as Safety of Care, but they have not examined this for the other measure groups.
- A TEP member asked what happens to the measure group weightings if a hospital decides not to report Safety of Care measures and focuses on Mortality instead, whether they would still be able to get a 5-Star, and what can be done to prevent loopholes like that.
    - Mr. Bagshaw clarified that hospitals do not get a choice of which individual measures to report. The measures included in the Overall Star Rating are required by at least one other CMS program; hospitals are required to collect and report underlying data for all measures and then CMS will determine if the hospital met the minimum case count for that measure and either publicly report or not publicly report the measure score.
    - Dr. Gettel added that if a hospital does not have enough case volume to report any Safety of Care measures, then weight of the Safety of Care measure group is redistributed across the other measure groups and the hospital could still get a 5-star rating if they meet the other requirements.
        - A TEP member asked about the repercussions of a hospital receiving 1-star and whether CORE has statistics that show the implications of using both reweighting and a Star Penalty as there is a drastic difference. They wondered if these approaches can be combined.
            - Mr. Bagshaw noted that applying both reweighting and a Star Penalty is an option under consideration and there is data included in the appendix of the slide deck that shows the results of applying both.
            - Dr. Gettel clarified these results are included on slide 61.
        - Dr. Venkatesh further clarified that hospitals cannot “choose” what measures to report as they report all measures for which they meet minimum case thresholds required by a reporting program — hospitals could theoretically choose what services to offer and that could impact their case volumes with respect to the thresholds but that has not been observed in practice.
- A TEP member requested confirmation that if CMS were to apply a Star Penalty, essentially the provider would be penalized twice, once for having a low score and then after the total score, they receive a second penalty for having a low Safety of Care score.
    - Dr. Gettel confirmed and discussed a scenario of a hospital with a Safety of Care score in the lowest quartile that received a 4-star rating despite good performance in other groups, and then after the Star Penalty is applied receives a final 3-star rating; he confirmed that one way of framing this is as a double penalty due to poor Safety of Care.
- A TEP member asked for clarification on the ultimate goal of re-tuning the methodology for the Star Rating, and whether the intent is to focus specifically on the poorest performers in Safety of Care, or

to achieve a broader sentinel effect across all hospitals regarding Safety of Care performance. They asked if we are trying to achieve both effects and which is most important.
    - Mr. Bagshaw noted the narrower goal is based on the observation that some hospitals can achieve a 5-star rating despite having a poor outcome for Safety of Care and therefore the goal is to avoid giving a "stamp of approval" to a hospital that cannot demonstrate a certain threshold of performance on Safety of Care. The broader goal is to increase the importance of Safety of Care within the Star Rating methodology to encourage safer care overall in alignment with CMS's goals.
    - A TEP member asked to what extent the CORE team is communicating with these lower performing hospitals or a subset of them to better understand the discrepancy in Safety of Care outcomes noting there may be just enough claims data to achieve a Safety of Care score. They noted the impact of Safety of Care on other measure groups such as Readmission or Complications for procedures such as elective joint revision; we are looking at this as a methodology problem to fix but they wanted to make sure there is a clear diagnosis of the problem first.
        - Dr. Gettel noted this was a good point, and clarified the CORE team has not communicated directly with these hospitals, but we can consider these contexts more. He agreed it was an important distinction whether this methodology adjustment would be intended to target the poorest performers in Safety of Care or to broadly emphasize Safety across all hospitals
- A TEP member noted concern about singling out Safety of Care and treating it differently than Mortality. They noted it seems inequitable and sends the wrong message when risk-adjusted mortality measures are among longest standing quality measures in the measure portfolio.
    - Dr. Gettel noted the CORE team ran similar analyses for Mortality and communicated similar concerns regarding the equal emphasis of Mortality and Safety in the current methodology; CMS requested prioritization of Safety of Care and the CORE team will share this input with CMS.
- A TEP member noted they were concerned about the focus on Safety of Care instead of the other measure groups. They added that smaller hospitals that do not have enough predicted events to calculate a Standard Infection Ratio (SIR) (for one or more HAI measures) are more likely to have fewer measures to report for Safety of Care. They elaborated although there is a minimum number of measures required, smaller hospitals' measure results could look very different from those of larger hospitals that do have sufficient volume, which is an important acknowledgement when considering reweighting and a Star Penalty based on only the quartile of performance.
    - Dr. Gettel noted the CORE team is currently running some analyses to evaluate the results based on hospital characteristics, including hospital size.
- A TEP member noted that some of the measures in the Star Rating methodology are required for two CMS payment systems and there are monetary disincentives for hospitals to not complete required data submission.
- A TEP member made another point inquiring whether any hospital should be able to achieve a 5-star rating if they are in the worst quartile of performance in any measure group. They understood the reasoning for looking at this for Safety of Care and noted it would seem to be fitting for the Mortality measure group and others as well.

- Another TEP member noted agreement with this comment and later shared similar feedback in the larger discussion.

Feedback from the Patient and Family Engagement Workgroup (PFE WG)

- Mr. Bagshaw noted the CORE team convened a PFE WG that consists of patients, family caregivers, and advocates. These options (reweighting, Star Rating Penalty/Cap) were discussed with them as well.
    - Mr. Bagshaw shared that the PFE WG had mixed reactions to CORE's presentation about reweighting with key feedback suggesting:
        - The initial weights were somewhat arbitrary or appear to be arbitrary;
        - There are not enough hospitals in this category to require a methodology change that would impact all hospitals;
        - The other measure groups are as important as Safety of Care and down-weighting them could potentially reduce incentives for hospitals to improve quality; and
        - Some participants expressed support for increasing the weight of Safety of Care to 30%, given the importance of Safety of Care and the modest reduction of weights for the other measure groups.
    - Mr. Bagshaw shared that the PFE WG also had the following mixed reactions to CORE's presentation about options for the Star Penalty/Cap:
        - The PFE WG tended to prefer the Star Penalty for 5-Star hospitals over a Star Cap;
        - There were concerns the Star Penalty/Cap option would not affect hospitals with insufficient Safety of Care information, even though Safety of Care might be a concern at these hospitals;
        - They also suggested all hospitals should be held accountable to the same standard, regardless of their initial Star Rating, and introducing a penalty could motivate them to improve their performance in Safety of Care to a greater degree than the current methodology.

TEP Discussion of Options for Reevaluating Safety in Star Ratings

- **Ms. Thottam asked the group the first discussion question:**
- **What are your thoughts on reweighting the Safety of Care measure group to have more influence than the other measure groups?**
- A TEP member shared their initial thought that reweighting the Safety of Care measure group to have more influence than the other measure groups solves the immediate problem, but wondered if in the future the Mortality measure group would warrant the same weighting adjustment. They added it is also important to consider the comments from the PFE WG that the weights seem somewhat arbitrary and could make it harder to determine if additional reweighting would be needed in the future.
- A TEP member liked the fact that reweighting would affect all hospitals as opposed to some of the other suggestions that are limited just to those hospitals with 5 stars. They thought safety is important for every hospital and liked the fact that reweighting is applied universally across the program.
- A TEP member stated that similar to what others stated, the reweighting is, and always will be a little bit arbitrary. They added that there has clearly been a decision made to make the measure groups equal, which makes sense for the Overall Star Rating. However, they thought if it is necessary to

overemphasize patient safety, then reweighting is a reasonable option. They also suggested reweighting is preferable because this option would allow hospitals excelling in Safety of Care to be rewarded for that performance even though they may be receiving lower scores in other measure groups rather than the Penalty or Cap options which only penalize hospitals.

- A TEP member agreed the weights seem arbitrary and noted that Mortality is equally important and only dealing with Safety of Care in the absence of the other measures may not be the most efficient option because additional reweighting might be needed in the future. They suggested that reweighting is a good idea but thought it needs to be considered more in the context of everything rather than just Safety of Care.
- A TEP member noted opposition to reweighting of Safety of Care stating that reweighting the Safety of Care score to 30% means the Mortality score is reweighted to 19.7% anrd it would be an over 10 percentage point difference between the weighting of Safety and Mortality. They did not think that was appropriate and opposed this.
- A TEP member noted opposition to reweighting Safety of Care for the various reasons offered by the group and described this option as a big miss because it comes across as trying to avoid lawsuits. Essentially, if a hospital is safe in the care they provide, they are going to reduce their liability and risk as a business. The TEP member added that it is important to focus on psychological or emotional safety in addition to physical safety (infections or readmissions). They did not agree with increasing Safety of Care to 30%.
- A TEP member noted it is hard to say which of Safety of Care or Mortality should be considered more important. They like the domains being equally weighted for a holistic picture of what people would expect a 5-star hospital to be. They asked about the underlying measures and wondered if this was a small numbers problem such that the Safety of Care measures do not correlate well with the other domains.
    - Mr. Bagshaw responded that small hospitals are more prone to see variation in their scores due to smaller sample sizes; with Safety of Care being a priority area for CMS, we would expect similar results. He elaborated that smaller sample sizes will tend to make some measure scores more variable and might subject smaller hospitals to greater variation (e.g., more or less likely to have a higher safety score based on chance) versus larger hospitals that would likely have a more stable estimate. Mr. Bagshaw noted that CORE focused on the Safety of Care measure group for this analysis but would expect similar results if other measure groups were examined. He highlighted that most hospitals that do well in a given measure group will tend to do better in the Star Rating overall, but there may still be outlier hospitals with discordant performance in one group compared to the others. Mr. Bagshaw noted that smaller hospitals might have fewer stable estimates over time and may be slightly more likely to have such discordant results, but the analyses presented today represent a snapshot in time for July 2023 results, therefore results could change year to year even if the methodology remains the same.
- A TEP member noted support for reweighting and highlighted how 233 hospitals received a lower Star Rating than they otherwise would. They noted they would prefer to have Mortality and Safety of Care both increased to an equal weight, and other measure group weights reduced.
    - Dr. Gettel noted the adjusted weights assigned to the measure groups were for the purposes of the analysis to see what increasing the weight of Safety of Care might look like. He noted reweighting of Mortality was a possibility if it is aligned with CMS goals.

- A TEP member stated they do not fully support reweighting at this time but noted a preference for reweighting compared to other options as reweighting is a methodological way to emphasize safety that impacts more than just a small subgroup of hospitals. They asked whether there is a significant correlation between Safety of Care and Readmissions and added that to the extent that some of the most egregious safety events are captured in readmissions already, they would like to see more analyses on that. If the goal is to emphasize Safety of Care across the board for hospitals, they asked then how this is communicated to consumers that may be making hospital preference choices. They suggested considering small focus groups of patients to understand how the patients perceive Safety of Care and if this feedback was aligned with CMS's priority, they might be in favor of emphasizing Safety of Care.
- A TEP member noted given that current set of measures that are included in the Safety of Care measure group, they were not in favor of reweighting to 30%, elaborating that the measures are not fully representative of "safety." While they acknowledged this was likely outside the scope of this group, they did not think a large enough population of patients in the hospital are represented for a few of the measure groups. They noted concern for the hospitals that do not end up getting scored on all eight of the safety measures. In their own system, they have a couple of hospitals that only routinely get scores on two of the measures and therefore, their concern is that with the Safety of Care weight increasing to 30%, if a hospital has just one infection measure it would get too much weight in their overall rating. Otherwise, they agreed with a lot of the other previously shared comments.
- A TEP member noted they did not support reweighting as proposed, one reason being arbitrariness; they thought the definition of patient safety was somewhat arbitrary in relation to the other measure groups. Second, they noted the other measure groups capture other elements of patient safety; for example, readmissions can be affected by lack of a safe environment. Third, they added the groups are constructed using z scores to standardize the scores which is appropriate to use when the data is normally distributed, but the distribution of many Safety of Care measures are notably skewed compared to measures in other measure groups. They stated this makes for a noisier grouping and maybe there is more misclassification in this group. For these reasons, they were apprehensive to give more weight to such a group with that unfortunate characteristic.
- A TEP member noted they were opposed to reweighting in part because of the diminishing of the value of the other categories and agreed with many others' remarks.
- A TEP member noted they would echo a lot of the other concerns that have been discussed already. They would prefer reweighting over the Star Penalty but are concerned about how it could impact hospitals' focus on the other domains. They suggested there is a balancing component between some of the other metrics and as was mentioned earlier between Mortality, Readmissions, and Patient Experience noting the example that when there are Safety of Care issues, it is going to impact Patient Experience. They suggested reweighting could unfairly reward hospitals that might have a good Safety of Care score based on only one measure in that measure group. They added that some measures in this domain may have very small sample sizes and some hospitals do not even provide services reflected by some safety measures at all and therefore would not have those metrics associated with their performance; if they perform well in one of the other infection metrics, they might be unfairly rewarded. Conversely in smaller facilities, one infection could make them look like they perform poorly. The TEP member acknowledged that one of the other points recently made is that Safety of Care is more than just the metrics that are represented in the category. The TEP

member noted there are a lot of other important factors building that safety foundation and culture of safety that are not represented in these options.

- A TEP member reiterated that the Safety of Care domain is not the only measure group that reflects patient safety and they thought that should be a consideration in trying to meet CMS's initiative to bring Safety of Care to the forefront. They appreciated the thought and consideration taken for these approaches and they were glad to see that we are not proposing the latent variable model (LVM) again. There is some additional consideration needed before proposing these changes.
- A TEP member agreed that the proposed reweighting is arbitrary. They elaborated that Safety is not necessarily more important than Mortality, and that it seems too messy to emphasize a specific measure group based on CMS priority at that specific point in time. If moving forward with a reweighting scenario, they suggested a formal process of an expert panel to reweight based on a diverse group's opinion. They preferred to keep the weighting the way it is.
- A TEP member noted there is significant interest in the new price transparency situations like variations in hospitals for the same procedure. They added that national question for them and their populations is whether higher priced hospitals offer higher quality, and therefore the 5-star rating is important. That said, they did not support reweighting as patients should not have to decide which is more important: a good surgical outcome or a safe place as every hospital should be safe, and the quality of outcomes is a separate issue.
- A TEP member agreed with previous remarks and considered how we can make it easier for patients to understand what they are going to get and how it will drive hospitals' behaviors in the preferred direction; they do not think that reweighting does that. If a patient sees a 5-star rating, they would likely assume that Mortality and Safety of Care, as well as the other three measure groups are highly scored. They stated it feels unfair for hospital to be receiving 5 stars when they may in fact be in the lowest performance category of one of the measure groups compared to other hospitals. They are unsure if increasing the Safety weighting would drive hospitals' behavior just by changing it.
- A TEP member originally thought increasing Safety of Care to 30% and reducing the weight of other measure groups seemed reasonable; after listening to the discussion, they are a little more concerned about what "Safety of Care" is. From a patient perspective, safety is a lot more than the eight measures currently in the group. Their preference is to keep the current weighting.
- A TEP member shared concerns that others have mentioned about trying to decide what is more appropriate for prioritizing weighting between Mortality and Safety of Care and they agreed they should be equally important. They shared the concerns mentioned by two other TEP members that many small hospitals only qualify for one or two measures, with one of those measures likely to be PSI-90; upside in that measure is limited because of the reliability weighting and therefore the best you can do is average. Another concern is for HAI measures for which the national average is trending down and may be impossible for small hospitals to meet if they have even a single event. For example, the national average C. diff rate is 0.4; if a small-volume hospital is predicted to have 1.2, then even a single infection puts their SIR at 0.8 (double the national average), but it may not be fair to say this hospital is truly in the bottom 25% for safety nationally.
- A TEP member noted they have not heard in the discussion how pervasive this issue is for 5-star performance. They asked when looking at historical 5-star performance, whether there is a group of 5-star hospitals that consistently have low quartile performance on Safety of Care and whether we see persistent low quartile performance for a subset of 5-Star hospitals in the other measure groups.

They were hesitant to recommend a new methodology for something that was examined only for one annual star release.
    - Dr. Venkatesh responded the persistence question is one CORE has not explored deeply, but CORE could—for example, look at whether there are 5-star hospitals below the first quartile in Safety of Care for this measurement period versus two consecutive measurement periods. He wondered if that would change anyone's perspective on whether and how to reflect this in the Star Rating.
- A TEP member noted agreement with prior comments and added that in general across these options of changing the current methodology, not specific to Safety of Care, they would probably support reweighting over a Star Cap because it keeps all the measures on the same comparative performance aligning with the intention of the Star Rating; however, like others have said, it should be specific to Safety of Care. They highlighted that in past TEPs, equal importance has been placed on Safety and Mortality and that when considering the calculation of these measures, and as others have pointed out, the Safety of Care measures can be quite volatile. They noted that the American Hospital Association released a white paper last year showing that it is especially the case for smaller rural hospitals and Critical Access Hospitals and therefore they did not support this method of reweighting.
- A TEP member noted that if the goal is to drive performance in these measure groups, then thought should be given to add in a workforce measure, particularly workforce stability and staffing which underlies performance on all of these measures. They acknowledged this was not the goal of today's discussion, but recommend it be considered.
    - Ms. Eve Rothenberg thanked them for this suggestion and noted the CORE team would discuss this and might reach back out with follow up questions after the meeting.
    - A TEP member offered to help if the CORE team decided to go in this direction as it is something important that is currently missing in measurement standards in healthcare; it would be a huge win for the workforce and patients.
- A TEP member asked about the overlap of poor Safety of Care performers in our dataset versus other CMS programs outside of the Star Rating (i.e., Hospital Acquired Condition (HAC) Reduction program). They wondered who the non-overlapping hospitals were, how many there were, and whether there is a benefit to highlighting those as poor performers.
    - Mr. Bagshaw noted CORE looked at the hospitals that would overlap between the penalty in the HAC program and the penalty in the Star Rating program and there was a very high overlap. He noted there is some misalignment because the Star Rating has a slightly different measure set and data period and because there are hospitals in the Star Rating that are not included in the HAC program. The CORE team considered the misalignment and ultimately focused on the Star Rating definition of "poor performance" because it is more inclusive of hospitals, depends only on data already within the Star Rating methodology, and does not require additional external data.
    - Dr. Venkatesh noted that only about half of all hospitals are in the HAC program — The Overall Star Rating is a distinct lever to promote quality improvement with a broader set of hospitals.
    - A TEP member asked if similar discussions about an increased emphasis on safety was also happening for the Hospital Value-Based Purchasing (HVBP) program.

- Ms. Thottam summarized that generally the group seems opposed to reweighting and acknowledged additional considerations the TEP flagged. Ms. Thottam asked Mr. Bagshaw and Dr. Gettel if they had reactions to the feedback shared so far by the TEP.
    - Mr. Bagshaw noted CORE's appreciation of the TEP's input and noted the mixed feedback that was generally unfavorable to reweighting; a main reason is sentiment that the measures themselves in the Safety of Care group were insufficient, either due to reliability concerns or just in terms of what they measure.
    - Dr. Gettel noted that as we have additional discussions about the Star Penalty and Star Cap that participants think about whether it is worthwhile to emphasize Safety of Care more than Mortality. He suggested the TEP think about sharing whether their comments regarding reweighting still apply for a Star Penalty and/or Star Cap. In comparing these two additional options, he wondered if some of the feedback that was shared previously still holds true, and if some of those points the group made are universal across all presented options.
    - **Ms. Thottam shared discussion questions for the Star Cap/Penalty, including:**
        - **What are your thoughts of an explicit Star Rating penalty or cap for poor safety performance?**
        - **Do you think an explicit Star Rating cap should apply only to hospitals in the worst-quartile of the Safety of Care measure group that would otherwise get 5 stars?**
        - **Or should an explicit penalty apply to hospitals of any star rating?**
- A TEP member questioned whether the Star Cap should only be applicable to Safety of Care or whether it should be applicable to any measure group, and whether it should be applied to the bottom quartile or to the bottom decile. They reiterated this idea goes back to the thought of keeping everything at the same comparative performance, rather than a specific measure group being targeted for this change. They asked if there might be consideration that if a hospital scores in the bottom quartile (or decile) of any measure group that the maximum star score they could achieve would be 4-Stars, rather than limiting this to one measure group.
- A TEP member agreed with the previous comment on applying the Star Rating Cap to all the measure groups. Between reweighting and the cap, they lean more toward the Star Cap, but with the same concerns that they had about the reweighting; a hospital should need at least three measures in a given measure group to have the Star Cap apply because those hospitals that have fewer measures are subject to more variability in their outcomes.
- A TEP member questioned how a Star Rating should be interpreted by a patient and noted that if something is going to be applied to some hospitals it should be applied to all hospitals, but that the Safety of Care group does not fully reflect the definition of "Safety of Care" they had in mind.
- A TEP member noted they were unsure, but slightly favored a Star Cap. They asked whether a hospital at the very bottom of the barrel for Mortality or Safety of Care, should even be able to get as high as a 4-star rating, and they did not know what this type of change would do to the rating system, but highlighted that 4 or 5 stars signifies an excellent hospital where high mortality rates and poor safety outcomes are not expected. They favor applying a penalty broadly and not just to that select group of hospitals.
- A TEP participant preferred a Star Rating Cap where only 5-star hospitals are affected. They stated penalizing all hospitals would result in a double penalty for those hospitals with limited measures and involvement, which are already getting penalized before that when their Star Rating is calculated. They noted that someone can be a great surgeon but if the recovery room and the post-

operative treatment is poor, it is really going to impact patients/hospitals in a negative fashion. Therefore, they expressed that Safety of Care needs to be the number one consideration, and particularly for those hospitals that want to identify themselves as having 5 stars; if hospitals do not do well in safety, they do not deserve 5 stars.
    - A TEP member noted agreement with this feedback.
    - Another TEP member agreed that adding more measures to the Safety of Care group would definitely address some of their concerns with this option.
- A TEP member noted the presented options seem a little too manipulative of the whole system. They understand there is an interest in Safety of Care, but it is being applied to one measure group without respect for the others. They noted this discussion is really about a small number of hospitals that fall within the "wrong" category, and to examine it only for Safety of Care rather than across all measure groups does not seem right. They suggested if a hospital is in the bottom decile of any of the measure groups, maybe they should not be able to achieve a 5-star rating, and that would be a more uniformly applicable process than singling out a single domain. They added that as measures in each group change over time, this problem will change as well; they thought a lot of progress had been made since the original methodology to get to a pretty good point.
- A TEP member agreed with previous comments and noted there is not a lot of difference between the options, and none seemed like the right approach. They did not have an opinion on either option because they thought more exploration is needed.
- A TEP member agreed with much that was already said. They noted the same concerns with this methodology, as with the reweighting, with small sample sizes and metrics not fully representing safety in hospitals, duplication of penalties with the same metrics being in multiple programs and reducing the importance of the other domains. They were drawn to the option of having a Star Cap and applying it to all the domains but also suggested considering another approach that could help elevate Safety of Care through other means, rather than some of these presented options which they felt seem a bit manipulative.
- A TEP member agreed that trying to single out this one measure group and applying some new techniques seems strange. They added if a change is going to be made, it should be applicable across all the measures and measure groups, therefore they do not strongly prefer one option over the other. They highlighted a need for consistency and confirmed they do not prefer reweighting, just because it cannot be applied consistently to all of the other measure groups.
- Mr. Bagshaw requested the remaining respondents consider whether their concerns are about the Safety of Care group as it is constructed and whether their opinion of the measure group would change if there were a broader set of Safety measures or whether the concern is based on the concept of emphasizing this measure group more than the others.
    - Dr. Gettel further asked: if the safety measures were more numerous or broader would that change the opinion regarding the safety domain being emphasized, or even if we had different measures would the TEP still oppose increasing the weight of safety?
    - Two TEP members asked what led to the focus on Safety of Care as opposed to the other measure groups. They acknowledged it is important but wondered why it is more important than the other measure groups.
        - Mr. Bagshaw clarified that CMS requested the CORE team examine opportunities to increase emphasis on safety based on the broader quality strategy and Universal Foundation.

- A TEP member noted that in order to be responsive to CMS's request to prioritize Safety, if they had to pick one of the options, they would prefer the Star Cap option, but they would change the parameters to increase the certainty of poor Safety of Care in two ways. They suggested that the majority of the measures, at least five of eight, must be reported for the Star Cap option to apply as fewer Safety measures would not be adequate data to determine whether a hospital performed poorly in Safety. They also noted they do not support using the lowest Safety of Care quartile to designate which hospitals would drop from 5 stars to 4 stars and that there should be more certainty, such a test of significance, to determine which hospitals are performing significantly worse in Safety compared to other hospitals.
- A TEP member agreed with the previous respondent and stated if they had to choose an option, they would also pick the Star Rating Cap so that a hospital is not able to receive 5-stars if they perform poorly in Safety of Care. They agreed that there needs to be a little more robustness around how these hospitals are identified. The TEP member expressed they were very influenced by what the current measures are in the program, especially when they thought about concerns specifically for the small hospitals. They added that even when looking at the options from the perspective of a hospital that reported all of the existing Safety measures, they would still only support the Star Cap option.
- A TEP member agreed with the previous few comments. They noted it depends on who those poor performers are – if they are already penalized by the by other programs, namely the HAC reduction program, they already have a strong incentive to improve Safety of Care performance and the Star Rating Penalty may be particularly punitive, but if this captures some hospitals that are not included or adequately disincentivized through the HAC reduction program they may support. They would like additional information on which hospitals would be affected to make a decision.
- A TEP member noted they favor the Star Penalty over to the Star Cap because 530 hospitals would receive a lower Star Rating compared to the Star Cap, in only 19 of the hospitals are captured. They were in favor of the penalty and were okay with penalizing all hospitals with any Star Rating.
- A TEP member agreed with previous remarks supporting the Star Cap or the Star Penalty over the reweighting. They wondered if the quartile cut point was selected to align with the HAC reduction program and if it was statistically significant at that cut point rather than going down to deciles for applying the policy. They elaborated by asking if using the quartile was the right place to make a meaningful distinction around underlying Safety of Care. Regarding the question of more and better measures, they recognized these are CBE-endorsed measures and so it is hard to say the measures are not good. If the Star Cap is not adopted for all the measure groups, they recommended considering in the future to apply the same cap to hospitals that are outliers in Mortality; they noted this could be a harder sell for Readmissions and Patient Experience, but applying the Cap to Mortality and Safety of Care would be foundational and might appear less arbitrary.
- A TEP member thought patients might be misled by seeing a 5-star rating for a hospital and would not expect it to be in the lowest category for Safety of Care. They favor either a Star Penalty or Star Cap and thought it should also be applied at the very least to Mortality in addition to Safety, but to be logically consistent probably to the other domains, as well. They noted in regard to the issue of quartiles or deciles, the general approach is called partitioning, and they generally do not favor partitioning. They suggested reconsidering the fundamental approach used to classify the 20 low performing Safety hospitals.

- A TEP member noted they agreed with the last speaker, and they would prefer the Star Cap over the Star Penalty, but they still were concerned with just focusing on Safety of Care and they thought CORE needs to look at Mortality as well before making a decision.
- A TEP member did not favor a Star Penalty or Star Cap for Safety of Care. They highlighted we are just conceptually looking at a 5-star rating as being like a de facto weight of one hundred percent for a small, select group of hospitals who maybe have one surgical site infection (SSI) in a year, and that essentially dooms them in the Safety of Care measure group. They noted from a methodological standpoint, it is a tough sell in terms of statistical significance and highlighted that to be a 5-star hospital, a facility could be in the 20th percentile of the Safety measure group, but statistically indistinguishable in that performance from a hospital that is in the 80th percentile of Safety of Care. They expressed that is completely arbitrary to be selecting that partition, and recommended basing it solely on statistical significance if applying a policy that penalizes a hospital's Star Rating.
- A TEP member noted they were not opposed to the Star Penalty or Star Cap. They noted they were unsure if a 4-star rating gives the strongest message for a consumer to understand that something is bad. They preferred a penalty applying to all hospitals and not just those in the 5-star category. They noted adding additional measures might make a difference because it would expand the currently narrow focus and there are many other things relative to Safety of Care, including those that cross over to other measures like Mortality and Patient Experience. They highlighted the importance of determining how to align this with other programs, like the HAC reduction program.
- A TEP member noted of all the options, they thought the Safety Cap for the worst quartile of hospitals, where they could only get up to 4 stars, makes the most sense if it is only these 20 hospitals of concern. They noted an example if they were building a regression model, and there were outliers, they would not redesign the entire model based on the outliers. They suggested what is fundamentally missing right now is an investigation into what is driving that relationship; if we accept that the system is generally working, then what is happening with these 20 hospitals to identify them as outliers. They thought from the consumer perspective there might be stronger messaging on Care Compare to indicate that these hospitals' star ratings were reduced or that there is a lack of reliability in certain areas. They expressed it might be a global change as well in terms of measure groups where there is only one measure right now, or something along those lines. They highlighted their support for the Star Rating using existing measures in the Safety of Care domain rather than adding more measures as it is practical, but they recommended that if measures were going to be added, it would be ideal to start from more of a theoretical framework of what should be in it to capture Safety holistically rather than adding on to the existing measures.
    - Mr. Bagshaw responded that it was hard to draw conclusions about what is going on at the 20 hospitals and which hospitals are performing poorly in Safety as there are no clear trends in hospital characteristics. He recommended reviewing the slides later in the deck with the characteristics of the poor performing hospitals, compared to the nation.
    - Ms. Rothenberg noted the hospital characteristics for the hospitals in the lowest quartile of Safety are listed out on slides 54 and 55.
- Ms. Thottam thanked the TEP for their input and noted the consensus seemed to lean towards the Star Cap option but echoed very similar concerns to those previously expressed regarding reweighting.

- Mr. Bagshaw appreciated the thoughtful responses and questions to take back for consideration. He expressed that of the options, it appears the TEP had a slight preference toward the Stars Cap, as it is the most limited in terms of the impact and would be applied to the outliers.

Conclusion, Closing Remarks, and Next Steps

- Ms. Thottam noted additional upcoming stakeholder engagement via a Provider Leadership Workgroup scheduled in November. CORE will then gather the stakeholder engagement input and share it with CMS. She noted the meeting summary would be shared with the TEP in early December.

Email Feedback Received Post-TEP Meeting

- At the conclusion of the meeting, CORE invited TEP members present to share any additional questions or feedback they may have by email. Additionally, CORE shared the meeting recording with TEP members who were unable to attend and similarly invited them to comment via email. Several TEP members subsequently emailed CORE to provide additional feedback.
- A TEP member noted that from the patient and caregiver perspective, the measures in the Safety of Care measure group do not seem like adequate indicators of safety and suggested adding more Safety measures might be more representative of the services provided and may make the information more relatable to consumers. They added if they had a serious illness, they would prefer to go to a university teaching hospitals with a 3-star rating rather than a community hospital with 5-stars. They were surprised that at the TEP meeting, no one mentioned the common concern about the sicker patient population at academic hospitals influencing the Star Rating that those hospitals receive, and they highlighted there is no easy solution to this but that it should not be ignored.
- A TEP member stated that reweighting is not the best option, especially if Safety of Care is the only measure group that is be increased in weighting; they noted they could support increasing Safety of Care and Mortality to 25%, adjusting Readmissions and Patient Experience at 20% and reducing Timely & Effective Care to 10%, but this still would not resolve the issue of the measures in the Safety of Care group.
- A TEP member noted that as the discussion proceeded, they more strongly supported the Penalty or Cap options, particularly the Star Cap option but they reiterated they would prefer to apply the Cap to both Safety and Mortality rather than only Safety. They questioned whether this would mean a hospital could potentially lose two stars. They also requested adding information on Care Compare to more clearly display information, such as the peer grouping (which reflects hospitals with more or less data available) and if a hospital has a given rating because it originally scored higher but got a penalty compared to a hospital with that initial rating that was not penalized.
- A TEP member recommended not reweighting or applying the Star Penalty or Star Cap in reaction to the concern of the 20 hospitals. They noted these options, particularly the Star Penalty and Cap, seemed random and to emphasize Safety per CMS's request, it could just as easily be recommended that the top performers in Safety should get a 5-star rating, or a 1-star increase regardless of how they perform in other areas. Therefore, instead, they recommended addressing this concern by adding a better link or clarification on Care Compare site to make it more intuitive for consumers to see measure group performance when considering the Overall Star Rating of a hospital. They highlighted that if a consumer could readily see that a 5-star hospital was "green" in all measure groups, but "red' in Safety, it might be more meaningful to use this method of showing each measure group with a red, yellow, or green star based on quartile.

- A TEP member stated the Safety of Care measures are crucial and supported reweighting to reflect their importance. They preferred the Star Penalty compared to the Star Cap because it can apply to any hospital regardless of their rating, and they expressed it could even be worthwhile to have a 2-star penalty for poor Safety performing hospitals as opposed to their rating only dropping by a single star.
- A TEP member supported the Star Cap over the other options but shared the same concerns with many other participants about the focus only on Safety and not Mortality as well. They suggested examining other cohorts based on the number of Safety measures that are reported as that may show less ability to earn a 5-star rating with bottom quartile in Safety.