AMERICAN MEDICAL REHABILITATION PROVIDERS ASSOCIATION

March 21, 2025

Submitted electronically via PQMsupport@battelle.org

Re: Appeal Submission for CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities

The American Medical Rehabilitation Providers Association (AMRPA) appreciates the opportunity to submit our **appeal of the endorsement decision for CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities.** AMRPA is the national trade association representing over 800 inpatient rehabilitation facilities (IRFs)[1], both freestanding inpatient rehabilitation hospitals and rehabilitation units of acute care general hospitals. AMRPA members are Medicare-participating providers with quality measure information publicly reported on the CMS Care Compare website. AMRPA has always looked to be a partner to regulating agencies and other key quality stakeholders in promoting meaningful and effective quality reporting in the IRF program, and we look forward to continuing this type of partnership with Battelle and the Partnership for Quality Measurement (PQM) moving forward.

AMRPA recognizes the importance of a consensus-based entity (CBE) and the Endorsement & Maintenance (E&M) process that "ensures measures submitted for endorsement are evidence based, scientifically sound, and both safe and effective, meaning use of the measure will increase the likelihood of desired health outcomes; will not increase the likelihood of unintended, adverse health outcomes; and is consistent with current professional knowledge." AMRPA believes that the PQM E&M process is essential to ensure that quality measures meet all endorsement criteria, including that measures are not administratively burdensome, provide the ability to distinguish high-quality care in and among IRFs, do not impact beneficiary access to IRF care, and result in better patient outcomes. AMRPA stands ready to work with the PQM to ensure that the E&M process includes all the information to ensure that endorsed measures meet all the criteria and deliver meaningful measures to IRF patients, providers, or policymakers.

[1] Inpatient rehabilitation facilities (IRFs) – both freestanding and units located within acute-care hospitals – are fully licensed hospitals that must meet Medicare Hospital Conditions of Participation (COPs) and provide hospital-level care to high acuity patients. IRFs' physician-led care, competencies, equipment and infection control protocols are just some of the features that distinguish the hospital-level care provided by IRFs from most other PAC providers.

Chris Lee · Chair, AMRPA Board of Directors
Vice President and Chief Operations Officer, Madonna Rehabilitation Hospitals
529 14th Street NW, Suite 1280, Washington, DC 20045 · Phone: 202-591-2469 · Fax: 202-591-2445

With these principles in mind, AMRPA is appealing the endorsement decision for CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities. As detailed in the E&M Guidebook:

If a measure's endorsement is being appealed, including an "Endorsed with Conditions" decision, the appeal must cite evidence that the appellant's interests are directly and materially affected by the measure, and the CBE's endorsement of the measure has had, or will have, an adverse effect on those interests. The appeal must also include one of three rationales:

- *Evidence exists that was available by the cycle's Intent to Submit deadline but was not considered by the E&M committee at the time of the endorsement decision and is reasonably likely to affect the outcome of the original endorsement decision.*
- *The CBE's measure evaluation criteria were not applied appropriately. The appellant must specify the evaluation criterion that they believe was misapplied and why.*
- *The CBE executed a procedural error (i.e., CBE's E&M process was not followed). The appellant must specify the error/process step, how it was misapplied/not followed properly, and how this resulted in the measure being endorsed.*

The AMRPA appeal will specify and address the following criteria required for eligibility:

1. AMRPA member hospitals' interests are directly and materially affected by the measure, and the CBE's endorsement of the measure will continue to have an adverse effect on those interests.
2. The CBE's measure evaluation criteria were not applied appropriately, as the criteria for Importance and Harmonization did not adequately identify, discuss, or resolve issues related to competing measures.
3. There were multiple procedural errors which resulted in the measure being erroneously endorsed.

AMRPA thanks Battelle and the PQM for the opportunity and consideration of our appeal of CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities. Please review the each of the following sections which detail our reasons for appeal and address the three appeal eligibility criteria mentioned above.

***

1. AMRPA member hospitals' interests are directly and materially affected by the measure, and the CBE's endorsement of the measure will continue to have an adverse effect on those interests.

As stated in the comments provided by AMRPA and some of the AMRPA member hospitals and hospital organizations, we believe that the implementation of and endorsement of this measure has (and will continue to) negatively impacted IRFs and their patients. There are three main issues that already have or will have an adverse effect on AMRPA member hospital interests:

A. IRFs must invest in technological tools and solutions to obtain the information necessary to manage performance of this measure, which have not been meaningfully considered from a cost and burden perspective.

The imputation method involved in this measure is and will continue to be overly difficult for clinicians to know the patient values and manage performance. CMS does not provide IRFs with real-time patient-level data on this measure, limiting the opportunity to identify factors impacting performance. Without a technological/software solution or other means to calculate the imputed values in real-time, clinicians will be left to wonder what the patient's function score will be at admission or discharge and would have to wait for CMS to publish measure data, well after a patient is discharged, to know whether their patient met or exceeded the expected discharge score. Under these circumstances, where the values are not readily available, performance on this measure becomes guesswork for any patient requiring the use of an imputed value. Quality measures are meant to be actionable and meaningful; however, the imputation method will make it extremely difficult for IRFs to know what is needed to improve performance and provide meaningful results to patients.

Additionally, while CMS and the measure developers have stated that "…this measure adds no additional provider burden since it would be calculated using data from the IRF-PAI that IRFs are already required to collect," IRFs in fact need to educate and train their clinicians on the new measure, incorporate discussion of this measure into their interdisciplinary team meetings, and create a solution that will calculate imputation values and the risk-adjusted expected discharge function score values in order to manage performance. Clinicians will need to understand which functional items are included in this discharge function score measure, the implications of the imputation method and the use of the CMS-defined "missing" codes, and how performance on this measure may differ from other quality measures. Clinicians will also need to understand how to review measure results on Provider Preview Reports and other CMS reporting tools. We would estimate that the time necessary to educate/train one clinician on this measure would take at least one hour, and the review of Provider Preview Reports and

other CMS quality measure reports to take at least one hour every quarter, or 4 hours annually. Using the adjusted hourly wages provided in the proposed rule, the cost of educating, training, and managing this measure for one clinician would be between $250 (Licensed Practical and Licensed Vocational Nurses) to $450 (Physical Therapists) annually. Combining these costs across the over 1,100 IRFs, this could represent, at a minimum, a total annual cost burden of between $275,000 and $495,000. Should additional education and training be provided to other clinicians, this will add another $50-90 per clinician per hour.

Technology-related costs should also be considered, as the imputation method and risk-adjusted expected scores require advanced calculations to be able to monitor patient progress toward their expectation. We do not speculate what the average cost for such software development may be but instead note that any costs associated with these needs have not been considered as part of the endorsement of this measure.

While the data elements utilized in the measure calculations are already part of the IRF-PAI, IRFs would require a significant investment in technology to implement the imputation methodology required to manage the performance of this measure. Providers that are unable to calculate imputed values on their own will have limited ability to identify what the patients' measures scores may be order determine what risk-adjusted target value is required to contribute towards a positive result. For these reasons, AMRPA believes that this measure does not meet the feasibility criteria required for endorsement, and the costs of managing this measure will require a diversion of resources that will negatively impact IRFs and their patients.

B. The public reporting of this measure can lead to negative unintended consequences and care disruptions, counter to the goals of Care Compare.

AMRPA members continue to be concerned about the suggestion that this is a "cross-setting" measure and the unintended consequences that may result from this designation. This measure differs meaningfully across post-acute care settings as the risk-adjustment methodology utilizes setting-specific covariates and setting-specific coefficients. For example, each setting has a model intercept for the expected discharge function score for each patient. For IRFs, the model intercept is 34.1701; for SNFs, the model intercept is 30.0118; and for LTCHs, the model intercept is 14.5276. This suggests that depending on setting, a patient with the same patient characteristics will need to meet or exceed different expectations, with IRF patients having the highest expectation and LTCH patients having the lowest expectation. To be considered a cross-setting measure, AMRPA members believe that the patient expectations should be consistent regardless of setting.

Additionally, in suggesting this as a cross-setting measure that is publicly reported, these measures have the unintended consequence of potentially being used to limit patient access to certain settings based upon results. While CMS and the measure developers documented concerns about providers denying access to certain patients who may not perform well on this measure, AMRPA questions whether consideration was given for referral sources using this information to direct patients to alternative settings which may not provide the appropriate services to produce high quality outcomes. We ask that the committee consider these unintended consequences when evaluating this measure for endorsement. This measure is publicly reported and the values publicly displayed on Care Compare can negatively influence care decisions - especially if this measure is used in discharge planning processes, referrals, or prior authorization determinations. These circumstances negatively impact all interested parties who may utilize this quality measures for making care determinations, and the unintended consequences will negatively impact IRFs and Medicare beneficiaries who would benefit from the IRF level of care.

C. This measure produces values that are inconsistent and are not representative of the functional improvements and discharge functional status that patients achieve as part of their IRF care.

As we will note in additional reasons for appeal below, AMRPA members believe that CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities is a competing measure for two existing endorsed measures, CBE #2635 Discharge Self-Care Score for Medical Rehabilitation Patients and CBE #2636 Discharge Mobility Score for Medical Rehabilitation Patients. The two existing endorsed measures separate patient function into two separate domains, Self-Care and Mobility, and utilize all the functional status items included on the IRF-PAI. Using separate measures and all the functional items for quality recognizes that IRF patients have significantly different functional abilities or disabilities requiring the intensive therapies provided only in the IRF setting. The newly endorsed measure combines these two domains of patient function, utilizing only 3 of the 7 Self-Care items and only 7 of the 17 Mobility items included in the functional assessment of IRF patients. Utilizing less than half of all functional items included in the IRF-PAI and existing endorsed measures produces values that are inconsistent and are not representative of the functional improvements and discharge functional status that patients achieve as part of their IRF stay.

Additionally, the public reporting of this measure on Care Compare shows that performance on the Discharge Function Score measures is inconsistent with the two other endorsed measures. For example, in the December 2024 Care Compare refresh data file, 693 of the over 1,100 IRFs with reported values (nearly 62%) appear to have a value for the Discharge Function Score measure that is less than their performance on both existing endorsed measures that utilize all functional items. Similarly, AMRPA

member IRFs are reporting that patients who meet expectations for the Discharge Self-Care and Discharge Mobility measures are identified as not meeting expectations for the Discharge Function Score measure.

Since the Discharge Function Score publicly reports performance that is lower than performance on existing endorsed functional measures that incorporate all functional items, this will negatively impact all interested parties who may utilize this quality measures for making care determinations, and the unintended consequences will negatively impact IRFs and Medicare beneficiaries who would benefit from the IRF level of care.

2. The CBE's measure evaluation criteria were not applied appropriately, as the criteria for Importance and Harmonization did not adequately identify, discuss, or resolve issues related to competing measures.

As detailed in the E&M Guidebook, Appendix D provides the PQM Measure Evaluation Rubric. It states when evaluating the Importance criteria, measure developers and reviewers should consider "why existing measures/quality improvement programs are insufficient for addressing this health care need."

The guidebook continues by listing that reviewers indicate that the Importance Criteria is "Not Met" if:

- *There is no description of other existing measures or programs or no search conducted to identify other existing measures or programs; OR*
- *Proposed measure has the same measure focus and target population as existing measure(s) and offers no advantage in terms of addressing disparities, feasibility, potential use, or scientific acceptability; OR*
- *Patient input does not support the conclusion that the measured outcome, process, or structure is meaningful or it does so with a low degree of certainty.*

The guidebook also provides the following context for reviewers to indicate that the Importance Criteria is "Met" if:

- *Description of existing measures or programs justifies the proposed measure's focus among the proposed measure's target population and/or the proposed measure is superior[6] to identified related or competing measures; AND*
- *Description of patient input supports the conclusion that the measured outcome, process, or structure is meaningful with at least moderate certainty.*

The footnote provided additional context related to what may be considered for determining that a measure is superior to identified related or competing measures:

[6] *Measure developers/stewards must document why the proposed measure is superior to any identified and/or competing measures and should include any literature used to support this position. For instance, clinical practice guidelines supporting the proposed measure do not support any existing measures identified; or the proposed measure's intentions vary across programs/payors, which requires the measure to be distinct from other existing measures; or the proposed measure captures a target population at higher risk such that the use of the proposed measure may close care gaps for a higher-risk population.*

Additionally, in the guidebook section on Harmonization, PQM provides the following:
The current health care quality landscape contains a proliferation of measures, including some that could be considered duplicative or overlapping and others that measure similar but nonidentical concepts and/or define patient populations differently. Such duplicative measures and/or those with similar but not identical specifications may increase data collection burden and create confusion or inaccuracy in interpreting performance results for those who implement and use performance measures. Resolving issues around harmonizing measures and handling competing measures is one of the key challenges. Developers/stewards must respond to the questions about harmonization in their measure submission.

These measure evaluation criteria were not applied appropriately as they relate to competing measures for the following reasons:

A. Measure developers did not provide a description of other existing measures and did not conduct a search to identify other existing measures as part of the measure submission.

On the PQM measure-specific website, under the Importance tab, the measure developers do not provide any information indicating that a search was conducted to identify other existing measures, nor did the measure submission include a description of existing or competing measures. It was not until the November 21st Listening Session that this omission was rectified, when AMRPA public comments identified the two competing measures that were not included as part of the measure submission. Even after the competing measures were identified, measure developers did not update measure specifications or include descriptions of other existing measures on the measure Importance criteria.

Based upon the PQM Measure Evaluation Rubric detailed above, this would have suggested that reviewers should have indicated that the Importance Criteria was "Not Met". However, the PQM E&M Staff Preliminary Assessment suggested that the Importance Criteria was "Met". Additionally, results from the Committee Independent Review were mixed, with 4 reviewers indicating that the Importance criteria was "Not

met but addressable", while 3 other reviewers indicated that the Importance criteria was "Met".

The disagreement in reviewer determinations show that the measure evaluation criteria were not applied appropriately and should not have proceeded to a vote of endorsement until such a time as the measure developer updated specifications to provide a description of other existing measures. Accordingly, we believe an appeal is appropriate and meets the eligibility criteria required for further evaluation of the endorsement of this measure.

B. Measure developers did not document why the proposed measure is superior to any identified and/or competing measures or include any literature used to support this position.

When competing measures are identified, for the Importance criteria to be met, the measure developer must provide documentation describing how the proposed measure is superior to the related or competing measures. In this case, measure developers did not provide any documentation to update the measure specifications related to Importance. The only documentation available that identifies the competing measures is in response to public comments. In the responses from the measure developer, no information or literature is provided to indicate that the proposed measure is superior to the competing measures. Information is provided to show that the Discharge Function Score is highly correlated with the two existing measures, but that does not indicate that the proposed measure is in any way superior, and in fact may indicate that the measure is duplicative without adding new value in measuring quality.

Instead of providing documentation or literature indicating that this measure is superior to existing measures, the measure developer indicated that this specific measure *"fulfills statutory requirements as specified in the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014. The IMPACT Act includes a mandate that CMS develop and report cross-setting measures. It specifically requires CMS to develop cross-setting quality measures that assess functional status and changes in function."*

While AMRPA understands the statutory requirements of the IMPACT Act of 2014, the statement above does not meet the Importance criteria for competing measures. Additionally, CMS could fulfill the statutory requirements with the existing endorsed measures.

Like in the prior section, based upon the PQM Measure Evaluation Rubric detailed above this should have indicated that the Importance Criteria was "Not Met". However, the PQM E&M Staff Preliminary Assessment suggested that the Importance Criteria was "Met". Additionally, results from the Committee Independent Review were mixed, with

4 reviewers indicating that the Importance criteria was “Not met but addressable”, while 3 other reviewers indicated that the Importance criteria was “Met”.

The disagreement in reviewer determinations show that the measure evaluation criteria were not applied appropriately and should not have proceeded to a vote of endorsement until such a time as the measure developer documented why the proposed measure is superior to any identified and/or competing measures or include any literature used to support this position. Accordingly, we believe an appeal is appropriate and meets the eligibility criteria required for further evaluation of the endorsement of this measure.

C. Measure developers did not address harmonization in their measure submission.

After identifying competing measures through public comment, the measure developers did not respond to how “*duplicative measures and/or those with similar but not identical specifications may increase data collection burden and create confusion or inaccuracy in interpreting performance results for those who implement and use performance measures.”.* As we detailed above and noted in public comments, this measure is already creating confusion in interpreting performance results for those who use performance measures in making care determinations.

Measure developers did not update the measure submission and only responded to public comments about competing measures. Those responses did not resolve any issues relating to harmonization or reducing confusion between different performance results.

Because the measure developers did not resolve issues with harmonization, the measure evaluation criteria were not applied appropriately and should not have proceeded to a vote of endorsement until such a time as the measure developer updated their submission. Accordingly, we believe an appeal is appropriate and meets the eligibility criteria required for further evaluation of the endorsement of this measure.

3. There were multiple procedural errors which resulted in the measure being erroneously endorsed.

The following procedural errors occurred that led to the measure being endorsed:

A. The measure should not have been referred to the Recommendation Group for endorsement without the Importance Criteria being met.

As detailed previously, because the measure evaluation criteria were not applied appropriately, there was an improper designation applied to the Importance Criteria on the Staff Preliminary Assessment. When the measure went to Committee Independent Review, the majority of reviewers indicated that the Importance criteria was "Not met but addressable". Following these reviews, concerns with the Importance criteria were not addressed, as the measure submission and specifications were not updated to reflect the identification of competing measures or provide any documentation to support that the proposed measure was superior to existing measures. The Advisory Group Feedback provided in the E&M Advanced Illness and Post-Acute Care Endorsement Meeting Discussion Guide further indicated a "Dissenting" opinion in the category of "Overlapping Measures," continuing concerns noted earlier in the process about competing measures that the measure developer never resolved.

Procedurally, the measure should have been identified as not meeting the Importance Criteria earlier in the process and required the measure developer to provide the information necessary to address the Importance Criteria ahead of the Recommendation Group consideration of endorsement. Instead, the measure moved forward in the process with the Recommendation Group advised that the Importance Criteria was met and that the competing measure issue was resolved. This procedural error led to the measure being endorsed, which meets the eligibility criteria for an appeal, requiring further evaluation of the endorsement of this measure.

B. The measure developers provided misleading responses to Public Comments and E&M Advisory/Recommendation Group queries.

During the E&M process, measure developers provided misleading responses to various questions about this measure. The examples below demonstrate inconsistencies in the information provided in response to questions about this measure:

- In response to questions about the number of comments received, it was suggested that there were only 12 comments, and, as such, the concerns were not representative of all IRFs. In fact, AMRPA comments are representative of over 800 IRFs, and the comments from other IRF providers would be

representative of hundreds of facilities. These comments should not have been considered as coming from only 12 IRFs, but instead representative of the overwhelming majority of IRFs.

- In response to consideration of this measure by a Technical Expert Panel (TEP), it was suggested that the TEP supported the use of the imputation method included in this measure. This statement is misleading, as detailed in the Technical Expert Panel (TEP) for Cross-Setting Function Measure Development January 26-27, 2022 Summary Report, published in April 2022.

  During the Recommendation Group discussion, in response to concerns about the imputation method, measure developers suggested that they had reviewed the imputation method with the TEP and obtained support in moving forward with this and the overall measure. There are several issues with the response provided by the measure developers. The measure and imputation methods reviewed by the TEP were based upon information and development conducted by Abt Associates. However, the measure considered for endorsement is developed by RTI, and differs significantly from the measure developed by Abt Associates that was reviewed by the TEP. The TEP reviewed multiple options for the imputation method, none of which are consistent with the methodology included in the current measure developed by RTI. Additionally, the TEP did not support a specific imputation method nor did it review the current imputation methodology. Instead, the TEP provided the following feedback:
  "Panelists asked clarifying questions about the ANA methods presented. One panelist wondered if patient characteristics were considered in the sampling method used to generate the gold standard for bias/MSE calculations. The PAC QRP Support team clarified that the gold standard was not a random sample. Instead, it was constructed by matching on function scores on assessed items. Propensity score stratification had also been tested. Another panelist mentioned changes in reimbursement occurring in 2019 and asked if ANA rates and related findings would be impacted by those changes. The PAC QRP Support team indicated that testing had been conducted in more recent data, and the findings were similar.

  In response to which ANA method to use for the cross-setting measure, panelists tended to favor statistical imputation with continued refinement to improve cross-setting performance. Panelists agreed that the current recode could be improved upon and reiterated that not all ANAs reflect dependence on a function activity. One panelist mentioned that rescale has some limitations that imputation does not, citing the incentive problem. Panelists tended to consider statistical imputation the most accurate approach to estimating missing values since it uses more information about the patient to impute scores. A few panelists expressed concerns about how the imputation method appears to be

performing in HH and expressed support for continued refinement. Panelists emphasized the importance of clear plain language to describe the method to providers but thought clinicians would prefer a method that is more accurate over one that is simpler."

- In response to questions about providers having the ability to access reporting on this measure and not require additional technological solutions, the measure developer indicated that CMS was providing reports such that there would be no reason to require any additional systems. This statement was also misleading, as the reporting that CMS provides only summarizes the measure result after a patient has been discharged and IRF-PAI information has been supplied to the CMS IRF database. CMS does not provide a reporting solution for real-time patient information that incorporates the use of the imputation method (when necessary) or the projection of the risk-adjusted expected discharge function score. These two calculations are necessary to set patient level goals and track functional improvement throughout the stay consistent with expectations for this measure. AMRPA believes that the inaccurate response from the measure developers suggesting that reporting was available influenced the decision for endorsement and did not indicate the actual reporting provided to address concerns about the feasibility of this measure.

We believe that these misleading statements should be considered a procedural error, as these statements were not independently reviewed or verified for accuracy. Without any opportunity for public comment, the Recommendation Group was not accurately informed and therefore the endorsement decision may have been made incorrectly. AMRPA believes that it is necessary to correct the record on these statements and re-evaluate the endorsement decision given the corrected information.

***

AMRPA thanks Battelle and the PQM for considering our appeal of CBE #4630 – Cross-Setting Discharge Function Score for Inpatient Rehabilitation Facilities to ensure that quality measures meet all endorsement criteria, including that measures are not administratively burdensome, provide the ability to distinguish high-quality care in and among IRFs, do not impact beneficiary access to IRF care, and result in better patient outcomes. AMRPA stands ready to work with the PQM to ensure that the E&M process includes all the information to ensure that endorsed measures meet all the criteria and deliver meaningful measures to IRF patients, providers, or policymakers.

Should you wish to discuss our appeal further, please contact Troy Hillman, AMRPA Director of Quality and Health Policy (thillman@amrpa.org / (202) 207-1129) or Kate Beller, JD, AMRPA President (kbeller@amrpa.org / 202-207-1132).

Sincerely,

Chris Lee
Chair, AMRPA Board of Directors
Vice President and Chief Operations Officer, Madonna Rehabilitation Hospitals

Karl Sandin, MD, MPH
Chief Clinical Officer, AMRPA
Chair, AMRPA Quality of Care Committee
Member, AMRPA Board of Directors
Medical Director, Arroyo Grande Hospital Rehabilitation unit
Assistant Professor of Surgery (Trauma), Creighton University School of Medicine