NATIONAL QUALITY FORUM
Driving measurable health improvements together

Surgery Fall 2021 Cycle: Public and Member Comments

Contents

Post-Evaluation Measure-Specific Comments on Surgery Fall 2021 Submissions

NQF #3639 Clinician-Level and Clinician Group-Level Total Hip Arthroplasty and/or Total Knee Arthroplasty (THA and TKA) Patient-Reported Outcome-Based Performance Measure (PRO-PM) (Recommended)

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 7954 (Submitted: 04/20/2022)

Council / Public: Health Professionals Council (HPR)

Level of Support: Member Does NOT Support

Comment

The American Medical Association (AMA) continues to have concerns with NQF #3639, Clinician-Level and Clinician Group-Level Total Hip and Total Knee Arthroplasty (THA/TKA) Patient-Reported Outcome-Based Performance Measures (PRO-PMs). While the AMA supports the assessment of patient-reported outcomes, we believe that the burden of data collection to the clinician, practice, and patient must be adequately addressed and the continued multi-step approach to risk adjustment must be reconsidered prior to implementation of this measure in the Merit-Based Incentive Payment System (MIPS).

Developer Response

The American Medical Association (AMA) submitted a public comment on April 20, 2022, for the National Quality Forum (NQF) Measure #3639, stating support for patient-reported outcome-based performance measure (PRO-PM) development but noting concerns about the burden of data collection and the multi-step approach to risk adjustment. We thank the American Medical Association for their comment and provide responses to their concerns below. During measure development we carefully considered the issue of burden and will continue to carefully consider burden for future measure implementation. The manner in which data collection for this measure is integrated into clinical workflows will be at the discretion of the clinician/clinician group, which allows them freedom and flexibility to choose an approach that meets their needs and considers their patients' needs. The Centers for Medicare and Medicaid Innovation's (CMMI's) Comprehensive Care for Joint Replacement (CJR) Model, the source of the data for measure development and testing of this PRO-PM, served as proof of concept for patient-reported outcome (PRO) data collection and submission.

Challenges to PRO data collection can be mitigated by strong leadership support, flexibility in rearranging clinical workflows to accommodate PRO data collection, ability to access PRO data in real-time for clinical decision making, and universal staff buy-in on the value of PROs in improving care and quality. This approach also serves to increase patient engagement; patients have expressed to us the importance of knowing what PRO survey results will be used for and noted a greater willingness to complete surveys if they are collected by their provider and used in shared decision making. This integration benefits patients by engaging them in discussions about potential outcomes of their surgery. In addition, during the development of the NQF endorsed hospital-level total hip and total knee arthroplasty (THA/TKA) PRO-PM (NQF #3559) on which this measure is based, the development team solicited extensive patient and provider feedback to ensure the

measure included low burden patient-reported outcome measure (PROM) instruments. Extensive input from patients has indicated strong support for a PRO-based performance measure following elective primary THA and TKA. The Centers for Medicare & Medicaid Services (CMS) will carefully consider these concerns and comments during future implementation planning. Regarding the multi-step approach to risk adjustment, the clinician- and clinician group-level THA/TKA PRO-PM uses the same risk model as the hospital-level THA/TKA PRO-PM.

The hospital-level THA/TKA PRO-PM development team used a consensus-based approach to identify and vet clinically relevant risk variables important in predicting the improvement outcome, including a systematic literature review and environmental scan, a survey of orthopedists, consultation with an expert clinical consultant, extensive input from the Technical Expert Panel (TEP), and detailed public comments. During development of the hospital-level measure, the TEP and measure development team felt that health literacy held particular relevance for a measure based upon PRO data and this variable is included in the risk model. In consideration of additional social risk factors, we recognize that patients with social risk factors may present later with more severity or greater comorbidity. The increased clinical risk this presents is addressed through the clinical risk factors in the model. When we evaluated the impact of social risk factors, it was in the context of what additional impact they may pose. Although neither dual eligibility, Agency for Healthcare Research and Quality (AHRQ) socioeconomic (SES) index lowest quartile, nor non-white race were statistically significantly associated with the measure outcome, they were included in the statistical approach to non-response adjustment of the measure due to their statistically significant association with survey response in measure development data and in the literature. We will continue to evaluate the relationship between social risk factors and the measure outcome and evaluate the risk model over time.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Standing Committee Response

Proposed Response: Thank you for your comment. The Standing Committee recognizes the importance of improving outcomes for individuals undergoing surgical procedures and the opportunity for clinicians to enhance their surgical practice using provider performance tools, such as patient-reported outcome-performance measures (PRO-PMs). The Standing Committee recognizes the commenters' concerns and discussed these topics extensively during the measure evaluation meeting. While the Standing Committee maintains the measure meets NQF criteria and should be endorsed, it urges the developer to continue to monitor these issues as the measure is implemented and to make updates as needed.

Ms. Tilithia McBride

Comment ID#: 8053 (Submitted: 04/29/2022)

Council / Public: Provider Organizations Council (PRO)

Level of Support: Member Does NOT Support

Comment

The Federation of American Hospitals (FAH) supports the development and implementation of patient-reported outcomes performance measures (PRO-PMs) but we also believe that additional

questions and work remain before their widespread use such as the degree to which multiple PRO-PMs could lead to survey fatigue for patients; the potential impact that additional PRO-PMs may have on the reporting of well-established measures such as HCAHPs and CG-CAHPS; the PRO survey must demonstrate that it assesses outcomes that are relevant to the clinicians and facilities being measured; the PRO survey and PRO-PM must be tested across a diverse set of patients and facilities; and the PRO-PM results demonstrate meaningful gaps in care on which quality improvement activities can be focused. The FAH requests clarification on which version of this measure is under NQF endorsement review. Specifically, there are the differences in the post-operative timeframe of 270-365 days that is included in this measure under endorsement consideration and the timeframe of 300-425 days included in MUC2021-107. We believe that the updated timeframe of 300-425 days is more appropriate given the feedback from the Technical Expert Panel on timing of the 1-year follow-up visit.

On review of the measure specifications, the FAH notes that multiple data points beyond the typical clinical variables are required to ensure that the measure results are adequately risk adjusted. The FAH supports the inclusion of these data points but we are concerned that the developer has not provided sufficient information on how these data are collected and what additional workload and time will be required. For example, several of the data elements needed for risk adjustment are derived from patient-reported surveys, which must be collected within 0-90 days pre-operative. No information was provided on the processes used by the clinicians and practices such as whether it required coordination with the hospital or if the burden of the additional data collection was placed on hospital staff on the day of surgery. To what extent did these requirements impact clinical workflows and were additional staff resources required? What additional costs might an individual clinician and practice encounter as a result of implementation of this PRO-PM? Alternatively, from the patient's perspective, did the additional questions seem relevant and was the point in time during which these additional data were collected appropriate? It would also be useful to understand whether there is a potential for individuals to prioritize the completion of one survey over another and therefore lead to negative unintended consequences on response rates for other PRO-PMs such as HCAHPS and CG-CAHPS? Furthermore, we question whether this measure has been adequately tested for reporting by clinicians and practices since the data used to assess reliability, feasibility, and usability are based on hospital implementation of these PRO-PMs in the Comprehensive Care for Joint Replacement Model.

The FAH is also concerned that the measure does not include a case minimum in order to achieve an acceptable minimum threshold for reliability. The FAH believes that the developer must increase the minimum sample size to at least 25 cases to produce a minimum reliability threshold of sufficient magnitude (e.g. 0.7 or higher). While the FAH strongly supports the inclusion of health literacy in the risk adjustment model, we believe that the risk adjustment approach used by many developers considers the identification and testing of social risk factors as supplementary to clinical risk factors. This approach was identified as a concern by the NQF Disparities Standing Committee. Given that this was a new measure, it provided an opportunity for CMS to include these factors within the testing of the model rather than the previous approach of "adding on" factors after the model is developed. This type of approach would assist clinicians, practices, and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation. As a result, the FAH believes that this measure lacks sufficient information on the

potential impact these social risk variables have on the risk adjustment model. The FAH does not support the endorsement of this measure at this time and asks that the Standing Committee re-consider their recommendation. Thank you for the opportunity to comment.

Developer Response

The Federation of American Hospitals (FAH) submitted a public comment on April 29, 2022 for the National Quality Forum (NQF) Measure #3639, stating support for patient-reported outcome-based performance measure (PRO-PM) development and implementation in general but noting concerns about the burden of data collection, survey fatigue, reliability, feasibility, and usability of the measure, timing of patient-reported outcome measure (PROM) and risk variable data collection, and approach to risk adjustment. We thank the Federation of American Hospitals for their comment and provide responses to their concerns below. This PRO-PM, based on the NQF endorsed hospital-level total hip and total knee arthroplasty (THA/TKA) PRO-PM (NQF# 3559), has been developed to assess improvement in pain and functional status for patients following an elective primary THA/TKA. Attribution to the clinicians performing the surgery was supported by the Technical Expert Panel (TEP), the expert clinical consultant, and the Patient Working Group members engaged in this measure's development. The PROM instruments chosen to measure the improvement outcome were supported for their importance to clinicians and patients, feasibility, low burden, ability to be used to inform care management decisions, and ability to inform healthcare quality improvement efforts. Patient-reported outcome (PRO) data collected at the hospital level in the Center for Medicare and Medicaid Innovation's (CMMI's) Comprehensive Care for Joint Replacement (CJR) Model were successfully linked to the clinicians conducting the THA and TKA procedures performed in these hospitals, which were selected to be representative of the case mix experienced at various hospitals nationwide. Variation in risk-standardized improvement rates for clinicians and clinician groups demonstrate meaningful differences in performance measure scores, with risk-standardized improvement rates ranging from 18.36% to 88.56% for clinicians and 20.86% to 85.90% for clinician groups. Measure score reliability across a range of minimum case volume thresholds was conducted for the Centers for Medicare & Medicaid Services' (CMS) consideration in future implementation planning.

We are confident in the reliability, feasibility, and usability of this measure based on these testing results and recommend continued assessment in reevaluation. During measure development we carefully considered the issue of burden and will continue to carefully consider burden for future measure implementation. The manner in which data collection for this measure is integrated into clinical workflows will be at the discretion of the clinician/clinician group, which allows them freedom and flexibility to choose an approach that meets their needs and considers their patients' needs. CMMI's CJR Model served as proof of concept for PRO data collection and submission. Challenges to PRO data collection can be mitigated by strong leadership support, flexibility in rearranging clinical workflows to accommodate PRO data collection, ability to access PRO data in real-time for clinical decision making, and universal staff buy-in on the value of PROs in improving care and quality. This approach also serves to increase patient engagement; patients have expressed to us the importance of knowing what PRO survey results will be used for and noted a greater willingness to complete surveys if they are collected by their provider and used in shared decision making. This integration benefits patients by engaging them in discussions about potential outcomes of their surgery. We do not anticipate that this measure will contribute to survey fatigue or negatively impact response to other measures such as Hospital Consumer Assessment of

Healthcare Providers and Systems (HCAHPS) or the Consumer Assessment of Healthcare Providers and Systems Clinician & Group Survey (CG-CAHPS). This PRO-PM's eligible population is procedure-specific, which reduces the likelihood of the same patients receiving the HCAHPS/CG-CAHPS and the PROMs for this measure. Additionally, there is no overlap in the data collection timelines for these measures; the HCAHPS or CG-CAHPS are typically administered two weeks after a hospital or clinician visit, months before the postoperative data collection for this PRO-PM.

This measure utilizes the risk variables finalized for the hospital-level THA/TKA PRO-PM on which this measure is based. The hospital-level THA/TKA PRO-PM development team used a consensus-based approach to identify and vet clinically relevant risk variables important in predicting the improvement outcome, including a systematic literature review and environmental scan, a survey of orthopedists, consultation with an expert clinical consultant, extensive input from the TEP (which included patient members), and detailed public comments. Patient Working Group members were also strongly supportive of the risk variables included in this measure. The timing of preoperative PROM and risk variable data collection for this measure, intentionally aligned with the hospital-level measure, was determined with extensive TEP and stakeholder input and chosen to give flexibility for providers collecting these data. The commenter's support for including health literacy in the risk model is noted. During development of the hospital-level measure, the TEP supported inclusion of health literacy due to particular relevance for a measure based upon PRO data. In consideration of additional social risk factors, we recognize that patients with social risk factors may present later with more severity or greater comorbidity. The increased clinical risk this presents is addressed through the clinical risk factors in the model. When we evaluated the impact of social risk factors, it was in the context of what additional impact they may pose. Although neither dual eligibility, Agency for Healthcare Research and Quality (AHRQ) socioeconomic (SES) index lowest quartile, nor non-white race were statistically significantly associated with the measure outcome, they were included in the statistical approach to non-response adjustment of the measure due to their statistically significant association with survey response in measure development data and in the literature.

We will continue to evaluate the relationship between social risk factors and the measure outcome and evaluate the risk model over time. Finally, we appreciate the commenter's support of a postoperative PROM data collection timeframe of 300-425 days after the procedure. The PRO postoperative data collection period finalized for CJR was 270 to 365 days after the procedure; these were the data used in the development and testing of this measure and in the NQF submission. However, we have heard from multiple clinical experts strongly recommending a refinement to the postoperative data collection period to better align with clinical workflow and typical one-year follow-up scheduling, and to allow for better postoperative PRO data capture. Based on extensive input, we have proposed measure specifications for future measure implementation with a postoperative PRO data collection period representing this small shift to 300 to 425 days after the procedure. We do not anticipate that this will impact improvement results; we do anticipate an increase in PRO response.

NQF Response

Thank you for your comment. NQF evaluates each measure as specified by the developer at submission. The Consensus Development Process (CDP) and Measures Application Process (MAP) are two distinct

review processes with the potential to review measures with specification variance. We encourage those who implement these measures to use the endorsed version.

NQF Standing Committee Response

Proposed Response: Thank you for your comment. The Standing Committee recognizes the importance of improving outcomes for individuals undergoing surgical procedures and the opportunity for clinicians to enhance their surgical practice using provider performance tools, such as patient-reported outcome-performance measures (PRO-PMs). The Standing Committee recognizes the commenters' concerns and discussed these topics extensively during the measure evaluation meeting. While the Standing Committee maintains the measure meets NQF criteria and should be endorsed, it urges the developer to continue to monitor these issues as the measure is implemented and to make updates as needed.

Public Comments on Surgery Fall 2021 Draft Report

N/A

Pre-Evaluation Measure-Specific Comments on Surgery Fall 2021 Submissions

NQF #3639 Clinician-Level and Clinician Group-Level Total Hip Arthroplasty and/or Total Knee Arthroplasty (THA and TKA) Patient-Reported Outcome-Based Performance Measure (PRO-PM) (Not Recommended)

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 7854 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on NQF #3639, Clinician-Level and Clinician Group-Level Total Hip and Total Knee Arthroplasty (THA/TKA) Patient-Reported Outcome-Based Performance Measures (PRO-PMs). The AMA supports the assessment of patient-reported outcomes but believes that the burden of data collection to the clinician, practice, and patient must be adequately addressed and the continued multi-step approach to risk adjustment must be reconsidered prior to implementation of this measure in the Merit-Based Incentive Payment System (MIPS). We also request clarification on which version of this measure is undergoing endorsement review since MUC2021-107 includes a different post-operative assessment timeframe. We believe that the alignment of the timeframe with the 1-year follow-up visit as recommended by the technical expert panel feedback per MUC2021-107 is preferable. On review of the measure specifications, we note that the information required for the numerator and risk variables includes multiple data elements from additional patient-reported surveys beyond those used to assess the patient-reported outcome of interest. Furthermore, this information is expected to be collected between 90 to 0 days prior to surgery. The AMA supports the inclusion of many of these variables within the risk model given their relevance to how patients may or may not be able to achieve improvement but questions whether CMS adequately assessed the feasibility and potential data collection burden to the clinician, practice, and patient, particularly since the

data used for measure development relies on hospital reporting through the Comprehensive Care for Joint Replacement Model.

The limited information on feasibility does not provide any detail on how the testing sites coordinated data collection across settings or on whom the responsibility of the additional items was placed. This question is particularly important since the specifications require clinicians and practices to collect data for one measure from 90 days pre-operatively to up to 425 days post-operatively, which the hospital is also likely collecting at the same time. The inclusion of this measure in addition to the one at the hospital-level further raises our concerns over how duplication of effort in collecting these data required for the measure numerator and risk adjustment variables can be avoided. The NQF submission form does not adequately address these concerns and the AMA urges CMS to complete additional testing around the feasibility of data collection and reduction of reporting burden prior to endorsement. Perhaps even more importantly, we would have expected to see an assessment from the patient's perspective on whether the timing and number of items solicited throughout this process were appropriate and does not result in survey fatigue, particularly now that they may have the hospital and clinician requesting the same data. For example, would the number of surveys throughout the pre-, intra-, and post-operative timeframes lead them to be less likely to complete other surveys such as HCAHPS or CG-CAHPS? CMS should also examine if whether the timing of data collection is appropriate such as if the pre-operative PRO-PM data were collected on the morning of the surgery, could stress and anxiety have impacted responses? We believe that it is critical to understand the potential impact and burden that could be experienced. While it may seem reasonable for one measure, if this measure is an example of how future measures could be specified, what is the potential long-term impact on patients, hospitals, clinicians, and practices as more and more PRO-PMs are implemented?

The AMA also believes that measures must meet ***minimum*** acceptable thresholds of 0.7 for reliability. We urge NQF to require the developer to set the case minimum at 25 cases in order to achieve this threshold. The AMA strongly supports the inclusion of health literacy in the risk model but remains concerned that CMS continues to test social risk factors ***after*** the assessment of clinical and demographic risk factors and it is unclear why this multi-step approach is preferable. On review of the Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors report, it is clear that the approaches to testing these data should be revised to strategies such as multi-level models or testing of social factors prior to clinical factors and that as access to new data becomes available, it may elucidate more differences that are unrelated to factors within a hospital's control. Additional testing that evaluates clinical and social risk factors at the same time or social prior to clinical variables rather than the current approach with clinical factors prioritized should be completed. The AMA believes that additional information on these concerns is needed prior to endorsement of this measure. We respectfully ask the Standing Committee to consider these comments and seek additional information prior to any decision on endorsement.

Reference: National Quality Forum. Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors. Final report. July 18, 2017. Available at: *https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85635*. Last accessed January 17, 2022.