Public Comment Summary Report

Project Title:

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

Dates:

The Call for Public Comment ran from Tuesday, October 19, 2021 to Wednesday, November 17, 2021.

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (YNHHSC/CORE) to develop a 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). The contract name is Development, Reevaluation, and Implementation of Outcome/Efficiency Measures for Hospitals and Eligible Clinicians, Option Period 2. The contract number is HHSM-75FCMC18D0042, Task Order Number HHSM-75CMC19F0001. As part of its measure development process, CORE requested interested parties to submit comments on the candidate or concept measures that may be suitable for this project. The goal of this measure is to assess the quality of care provided to Medicare beneficiaries by clinicians and clinician groups performing elective primary THA and/or TKA procedures using patient-reported outcome (PRO) data. CORE requested interested parties to submit comments on the measure specifications, reporting at the clinician-level or clinician group-level, and future measure implementation.

Information About the Comments Received:

The measure developer solicited public comments by email notification to CMS listserv groups, email to relevant stakeholders and stakeholder organizations, and posting on the CMS Public Comment website.

We received five responses on this topic. Specifically, from:

- One hospital/health system (Henry Ford Health System)
- Four medical associations and societies (American Physical Therapy Association [APTA], American Association of Hip and Knee Surgeons [AAHKS], Federation of American Hospitals [FAH], and the American Medical Association [AMA])

Stakeholder Comments

General Stakeholder Comments:

We received comments on various aspects of the measure. Comments focused on the measure methodology, including the cohort, outcome, reliability and validity, risk adjustment, approach to non-response, attribution, and measure calculation; and on the feasibility of implementing and using the measure while reducing burden.

Most commenters were supportive of the development of a clinician- and clinician group-level quality measure that uses PRO data following elective primary THA and/or TKA procedures. Four of the five commenters expressed support for the proposed measure and its potential impact on health outcomes and quality improvement. However, four commenters conveyed concern related to adjustment for social risk factors, and five commenters expressed concern about implementation, including data collection and data submission burden, possible cherry picking, and possible duplication of efforts related to measure implementation across programs/settings. For more details about the commenters' concerns, see the summaries of the comments received and our responses below.

Measure-Specific Stakeholder Comments:

General

There were two general comments in support of the use of PROs.

- One commenter expressed general support for the assessment of PROs, and another commenter stated support for the development and implementation of PRO-PMs.

Response: We appreciate the commenters' support for development and implementation of PRO-PMs. We agree that assessing PROs is valuable and aligns with patient-centered measurement.

There were two comments about the use of PROs for the measurement of THA and TKA and functional status following these procedures.

- One commenter agreed that individuals undergoing THA and TKA are ideal candidates for assessing PROs. Another commenter expressed support for the use of PROs as the best available means for patient-centered measurement of functional status improvement.

Response: We appreciate the commenters' support of the measure and the PRO approach to assessing improvement following THA and TKA.

There was one comment about the re-specification of the hospital-level THA/TKA PRO-PM.

- The commenter was supportive of the re-specification of the hospital-level THA/TKA PRO-PM as a clinician- and clinician group-level measure for the Quality Payment Program (QPP).

Response: We appreciate the commenter's support for this measure.

Cohort

One commenter addressed expanding the cohort to include THA/TKAs performed in the outpatient and ambulatory surgical center (ASC) settings.

- The commenter expressed support for testing the measure in the outpatient and ASC settings in the future given the expanding volume of THA/TKAs in those settings.

Response: We appreciate the commenter's feedback. We agree with the importance of evaluating THA/TKA procedures in the outpatient and ASC settings. CMS will evaluate possible expansion to the applicable settings during implementation planning.

Outcome: Patient-Reported Outcome Measures (PROMs)

One commenter noted support for the patient-reported outcome measures (PROMs) included in the measure.

- The commenter appreciated the in-depth literature review pertaining to the PROMs for this specific population. They supported use of the PROMs (Hip dysfunction and Osteoarthritis Outcome Score for Joint Replacement [HOOS, JR], Knee injury and Osteoarthritis Outcome Score for Joint Replacement [KOOS, JR], Veterans RAND 12 Item Health Survey [VR-12], Patient-Reported Outcomes Measurement Information System [PROMIS]-Mental Health), as well as the Oswestry Disability Index question on back pain, and the question pertaining to pain in the non-operative joint.

Response: We appreciate the commenter's feedback regarding our literature review and support for using the included PROMs and risk variables.

Outcome: Timing of PROM and risk variable data collection

One commenter addressed the expanded postoperative PROM data collection timeframe.

- The commenter supported the recommendation to expand the postoperative PROM data collection timeframe from 270-365 days to 300-425 days as this better aligns with postsurgical appointment scheduling.

Response: We appreciate that the commenter agrees with the decision to update the postoperative PROM data collection timeframe and its alignment with clinical practice. Clinician experts and patients included in our Technical Expert Panel (TEP), Patient Working Group, and Clinical Working Group also strongly supported this decision.

Two commenters addressed the timing of preoperative PROM and risk variable data collection.

- Both commenters expressed concern about the 90-0 day preoperative data collection window and whether the preoperative data are collected at the appropriate point in time, particularly if collected the morning of the procedure. One commenter expressed concern about the potential stress and burden on patients from whom data are collected the morning of the procedure and questioned whether that could impact responses. Another commenter questioned whether staff would be burdened by collection of data the morning of the procedure.

Response: We appreciate the commenters' feedback. The timing of preoperative PROM and risk variable data collection for this measure is intentionally aligned with the hospital-level THA/TKA PRO-PM. The hospital-level THA/TKA PRO-PM development team engaged stakeholders to determine appropriate timing of PROM and risk variable data collection, who supported the preoperative PROM and risk variable data collection timing noting it would give flexibility for providers collecting these data. As clinicians can and do use PRO data to inform shared decision making with patients regarding the decision to proceed with THA and/or TKA procedures, we anticipate clinicians and clinician groups will have access to PRO data prior to surgery.

Outcome: Substantial Clinical Benefit (SCB)

One commenter addressed the substantial clinical benefit (SCB) thresholds.

- The commenter agreed with the measure's use of the SCB thresholds identified by Stephen Lyman and colleagues.[1]

Response: We appreciate the commenter's feedback and support for the use of the SCB thresholds for measuring clinical improvement following THA and TKA procedures.

Risk Adjustment

Two comments addressed the variables required for risk adjustment.

- Both commenters expressed support for the inclusion of the identified risk variables as they are relevant to how patients may or may not be able to achieve improvement and ensure the measure results are adequately risk adjusted.

Response: We appreciate the commenters' feedback and support for the risk variables included.

Three comments addressed the approach to social risk factors.

- Two commenters expressed support for the consideration given to social risk factors in the risk-adjustment model.

Response: We appreciate the commenters' support for this measure's approach to social risk adjustment.

- One commenter supported the risk model developed for this measure and recommended continued examination into inclusion of sociodemographic factors. Specifically, they recommended ongoing evaluation of the risk model to ensure the risk factors identified from patients within the Comprehensive Care for Joint Replacement (CJR) dataset are appropriate when applied to the entire Medicare population. The commenter also recommended stratification by proportions of dual-eligibility similar to the stratification approach used in the payment determination for the Hospital Readmissions Reduction Program.

Response: We appreciate the commenter's support for the risk model. 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. As this measure assesses patients undergoing an elective procedure where known disparities exist, we will continue to assess the impact of social risk for this measure, including the risk model, moving forward. We will share the feedback to consider stratification by proportions of dual eligibility for future measure implementation with CMS for potential future implementation planning.

There were three comments that addressed health literacy and other disparities.

- Two commenters supported the inclusion of health literacy in the risk model. However, the commenters expressed concern with the risk-adjustment approach, which considered other social risk factors as supplementary to clinical risk factors by assessing them after clinical and demographic risk factors. The two commenters recommended ongoing evaluation of the risk model, specifically consideration of clinical and social risk factors at the same time or social risk factors prior to clinical variables.

Response: We appreciate the commenters' support for including health literacy in the risk model, which was identified by clinical experts and stakeholders as an important predictor and is associated with a range of social risk factors. During development of the hospital-level THA/TKA PRO-PM, the TEP and measure development team felt that health literacy held particular relevance for a measure based upon PRO data. 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, and detailed public comments. The presence of social risk factors often leads patients to present later, more severely, or with 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 in addition to clinical risk factors; therefore, we considered social risk factors in this sequenced approach. We will continue to evaluate the relationship between social risk factors and the measure outcome and evaluate the risk model over time.

- One commenter expressed concern regarding the inconsistent health literacy findings in the testing of the risk model and the adequacy in which this domain is understood to impact PROs. Given the limited research pertaining to health literacy levels, patient engagement, system-level literacy, and PRO scores, the commenter recommended further investigating the impact these disparities may have on outcomes for THA/TKA patients.

Response: We appreciate the commenter's feedback and suggestion to further investigate the impact of health literacy on the measure outcome moving forward. We note the varying impact of health literacy in the risk model over multiple years and intend to continue to evaluate this factor. We understand that patient engagement and system-level literacy are related to health literacy and have an impact on PROs. We will continue to assess the impact of health literacy, and consider other potential risk factors available for assessment, with the measure outcome over time.

Reliability/Validity

One comment addressed the reliability and validity approach and results.

- The commenter expressed support for the examination of reliability and validity at the data element and measure score levels. They noted the results speak to the ability to identify meaningful differences of the proposed measure.

Response: We appreciate the commenter's support for the approach to and results of examining reliability and validity of this measure.

Non-Response Bias

Two commenters addressed non-response bias.

- One commenter expressed support for the measure's approach to addressing non-response bias. One commenter recommended continued considerations for addressing non-response bias.

Response: We appreciate these comments and agree with the importance of addressing potential non-response bias for this measure. We will continue to account for potential non-response bias using stabilized inverse probability weighting during future measure implementation and will continue to evaluate response rates in the future.

One commenter addressed the wide variation in procedure volume across providers in the context of response rates.

- The commenter recommended CORE consider the wide variation in volume across different locations or settings, such as rural versus urban versus suburban, as well as clinician versus clinician group.

Response: We appreciate the commenter's suggestion and will monitor the relationship between procedure volume and response rate. CMS commonly requires a minimum case volume for public reporting of measure results and we will share this with CMS for consideration in potential future implementation planning.

Attribution

Two comments supported the attribution approach.

- Two commenters supported and appreciated the modifications that were made to this measure's attribution approach. One of the commenters urged CMS to consider adopting this approach for other quality and cost measures in the Merit-based Incentive Payment System (MIPS).

Response: We appreciate the commenters' support for this measure's attribution approach.

Two comments addressed attribution of the measure to clinicians and/or clinician groups.

- Both commenters expressed their preference for having the measure reported at both the clinician- and clinician group-levels. One commenter expressed this approach would be preferable given the significant changes in practice that take place when individual clinician practices combine into larger groups or individual clinicians join larger groups. The other commenter expressed this approach would help to maximize participation in the measure.

Response: We appreciate the commenters' feedback. CMS will carefully consider this feedback during potential future implementation planning.

Burden

One commenter addressed the relevance of the additional questions that are required for risk adjustment.

- The commenter expressed concern, from a patient's perspective, regarding the relevance of the multiple data points that are required for risk adjustment beyond the typical clinical variables.

Response: We thank the commenter for this feedback. We understand the importance of accounting for patient-level factors that are clinically relevant and have strong relationships with the outcome as well as minimizing burden of collection of variables from the patient and provider perspectives. The MIPS THA/TKA PRO-PM utilizes the risk variables finalized for the National Quality Forum (NQF)-endorsed (NQF#3559) 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 TEP (which included patient members), and detailed public comments. Patient Working Group members engaged by the MIPS THA/TKA PRO-PM development team were also strongly supportive of the risk variables included in this measure.

Three commenters addressed data collection and reporting burden for clinicians and clinician groups.

- All three of the commenters expressed concern regarding data collection and the potential impact on clinical workflows, the potential requirement for additional staff resources, and costs associated with implementation of this PRO-PM. Two of the commenters urged for additional testing around the feasibility of data collection and the reduction of reporting burden prior to implementing this measure. One of the commenters recommended delaying inclusion of this

measure in the QPP for two years to monitor and assess its use at the hospital-level before expanding to the clinician level.

Response: We understand that PRO data collection and submission are inherently more resource consuming than CMS's claims-based quality measures. Although PROMs are not universally collected prior to and following THA and TKA procedures, incentivized PRO data collection within CMS's CJR model presents a proof of concept for feasible, low burden collection of PROs for quality measurement. Challenges to PRO 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. 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. In addition, the hospital-level THA/TKA PRO-PM development team solicited extensive patient and provider feedback during development to ensure the measure included low burden PROMs. While PRO-PMs require providers to integrate data collection into clinical workflows, this integration provides opportunity for PROs to inform clinical decision making and benefits patients by engaging them in discussions about potential outcomes. CMS will carefully consider these concerns and comments during potential future implementation planning.

Two comments addressed survey fatigue for patients.

- Both commenters expressed concern regarding survey fatigue and whether the number of surveys may cause patients to prioritize surveys over another, which could lead to negative unintended consequences on response rates for other measures which use patient-reported data, such as Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

Response: We appreciate the commenters' feedback. We designed the measure to illuminate a patient's pain and functional status before and after a THA or TKA, which is different than other surveys, such as HCAHPS, that captures patient experience. With regard to the comment that the MIPS THA/TKA PRO-PM may have a reporting impact on other measures, such as HCAHPS, we anticipate a minimal impact to other measures as the MIPS THA/TKA PRO-PM's eligible population is procedure-specific which reduces the likelihood of the same patient receiving the HCAHPS and PROMs. Additionally, the MIPS THA/TKA PRO-PM preoperative assessment (90 to 0 days before procedure) and postoperative assessment (300 to 425 days following procedure) timeframe is different than HCAHPS, which is two weeks after a hospital visit.

Measure Implementation: Response Rates

One commenter addressed minimum response rate targets.

- The commenter expressed concern at the target for minimum response rates. Given the incentive provided in CJR for collection targets being met, the commenter is concerned that the CJR experience is not appropriate for determining collection rate targets for national reporting. They recommended voluntary reporting of a structural measure to assess the viability of PROMs collection by clinicians and clinician groups, which results in gaining quality points. These efforts could then be used as a means of establishing the expected response rate. In addition, they suggested response rates start at low levels to provide a minimum basis to establish the expected response rate, which could then be raised to find any potential ceiling.

Response: We appreciate the commenter's feedback. CMS will consider this feedback during potential future implementation planning.

Measure Implementation: Data Collection

Three comments asked for clarification regarding the process and responsibility of collecting the PROMs required for this measure.

- Three commenters expressed concern that information was not provided on how data for this measure would be collected and who would be responsible for collecting preoperative PROM and risk variable data and postoperative PROM data. Two of the commenters specifically requested information on how data collection efforts would be coordinated between providers and hospitals.

Response: We appreciate the commenters' feedback. The measure is designed to allow for flexible data collection for clinicians and clinician groups. The data for measure development and testing was collected as part of the CJR model; CJR is a hospital-level model and does not specify mode or approach to PROM data collection. Clinicians and clinician groups currently collecting PROM data use a range of methods, including paper, telephone, and electronic formats. During conversations with providers and patients about PRO-PM implementation, stakeholders noted a preference to allow for multiple collection modes to adapt to needs of patient populations (e.g., a patient without access to a computer would not be able to fill out a survey only sent via email but may be able to respond if the survey was collected in person). Regarding the entity responsible for inviting patients to complete the PROM and risk variable data preoperatively and the PROM data postoperatively, we recognize clinicians and clinician groups will need to determine a PROM data collection approach that will work for their resources and workflows. In addition, we recognize the importance of allowing clinicians, practices, and hospitals flexibility to determine a PROM data collection approach that supports alignment and avoids duplication.

One commenter addressed incentives for the collection of PROMs.

- The commenter recommends an incentive for the collection of PROMs. They noted that an incentive program providing additional quality points for high performances could encourage clinicians and clinician groups to improve systems for PROM data collection and submission, which could make the measure more meaningful as response rates improve. They also recommended the development of a new evaluation and management (E/M) code to support the collection of PROMs, which could in turn support clinicians and clinician groups in developing infrastructure for data collection and support the participation in registries for data submission.

Response: We appreciate the commenter's feedback. CMS will consider this feedback during potential future implementation planning.

One commenter addressed the issue of cherry picking.

- The commenter expressed concern that clinicians or clinician groups may cherry pick survey respondents by choosing those most likely to report a good outcome and not inviting others to complete the surveys. They wondered whether the measure results and risk model would perform as well, particularly in terms of possible cherry picking during possible future measure implementation.

Response: We understand that cherry picking is a possible issue for PRO-PMs, which rely on voluntary submission of data from patients and consequently may not require 100% submission rates, and we appreciate the commenter's concern. Our approach to account for potential non-response bias will to some extent mitigate the impact of cherry picking survey respondents. Additional approaches to reduce the potential for selection bias include submission of preoperative

data prior to and separate from the submission of postoperative data and monitoring response rates. CMS will consider this feedback during potential future implementation planning.

Measure Implementation: Data Submission

There were two comments regarding data submission.

- Both commenters recommended and supported the use of clinical registries for the transmission of PROM data. One of the commenters noted that the use of a clinical registry for PROM data submission would be possible and cost-effective. The other commenter noted that the use of a clinical registry submission would encourage provider utilization of Qualified Clinical Data Registries (QCDRs).

Response: We appreciate the commenters' suggestions regarding the use of a clinical registry to transmit PROM data. CMS will consider this feedback during potential future implementation planning and aims to ensure the measure is implemented in an efficient and cost-effective way.

Measure Implementation: Measure Reporting Across Programs and Settings

Two commenters addressed the potential for duplicate measure reporting.

- Both commenters expressed concern that if the measure was implemented in MIPS, clinicians and clinician groups would be reporting on the same data as the hospital-level measure. They noted that there was no identification of ways to mitigate potential duplication in reporting across programs.

Response: We appreciate the commenters' feedback. The Meaningful Measures Framework supports alignment of quality measures across programs to minimize provider burden.[2] CMS will consider this feedback during potential future implementation planning.

Usability

One commenter addressed how the data will be reported and presented.

- The commenter expressed the importance in reporting and presenting the data in meaningful ways in order to enable beneficiaries to make educated decisions about their care, and to enable clinicians to reflect and improve on the quality of care that they provide.

Response: We appreciate the commenter's feedback and agree with the importance of presenting the measure results in meaningful ways to support patients in making decisions about their care. CMS will work to provide measure results that are reported and presented in meaningful and useful ways to patients.

Other

One commenter addressed the possible correlation between PROM results and setting and ancillary service utilization.

- One commenter suggested investigation into the correlation between PROM results, setting of the procedure, and ancillary service utilization. The commenter also asked for consideration of how various quality measures correlate for broader implications for improvements in access, appropriate utilization and cost, and overall PROs.

Response: We appreciate the commenter's feedback and will consider this recommendation in future measure reevaluation activities as feasible.

Preliminary Recommendations

We plan to incorporate the recommendations received during public comment into the development and future implementation of our measure. Specifically:

- We will continue to assess the impact of social risk during future measure reevaluation.
- We will continue to evaluate the risk model, including social risk factors, during future measure reevaluation.
- We will continue to monitor response rates and the relationship to provider volume during future measure reevaluation.
- We will ensure CMS is aware of the public input received for any future implementation planning, including approaches to setting minimum case volume thresholds for data collection, incentivization of voluntary reporting, and the use of clinician registries for data submission.

Overall Analysis of the Comments and Recommendations

The feedback on the measure and the measure's proposed use for clinician- and clinician group-level reporting was positive. Commenters identified several technical issues and concerns related to measure implementation that we will address through future measure implementation planning and reevaluation. TEP members reviewed the public comments and supported the recommendations, particularly noting the importance of incentivization of PRO data collection.

Public Comment Verbatim Report

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
10/20/21	David R Nerenz, PhD, Henry Ford Health System	Hospital/Health System	MIPS THA/TKA PRO-PM	In response to the invitation to comment on: 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). The development team has done an exemplary job of examining reliability and validity at both data element and measure score levels, and in developing and testing risk-adjustment models and adjusting for non-response bias. The attention to social risk adjustment is careful and thoughtful. The presentation is clear and straightforward and the results speak to the reliability, validity, and ability to identify meaningful differences of the proposed measures. Just two comments/suggestions on the analyses and presentation: The analysis of non-response bias is very carefully done, but seems to assume that all eligible patients will at least be asked to complete both baseline and follow-up surveys. The description in the technical report does not indicate which entity was doing this invitation in the data set used for testing - was it the surgeon, the surgeon's practice, the hospital, an independent survey group under contract to CMS or CMMI, or someone else? In future use of the measure, a surgeon or group being evaluated may be able to "cherry-pick" survey respondents by choosing those most likely to report a good outcome and not inviting others to complete the surveys, unless there is an audit function that would detect this behavior. The risk-adjustment model is modest in terms of goodness-of-fit, suggesting that there is still some room for identification and use of factors predictive of outcomes to "cherry-pick" patients that are not included in the risk-adjustment model. The analysis of non-response bias does not need to be altered, but the discussion of use of the measure may want to discuss the problem of "cherry-picking" and discuss how properties of the model might change if taken outside of the context of the specific demonstration projects whose data were used for testing. If used in MIPS, for example, by surgeons and groups not participating in a CMS or CMMI demonstration, would the measure be as reliable and valid and otherwise perform as well, particularly in terms of a possible "cherry-picking" concern? Then, a more minor point. Measure score reliability seems to be impressively high (page

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
10/20/21	Continued from the previous page	Continued from the previous page	Continued from the previous page	49). Given that, the reference to Landis and Koch (page 50) is unnecessary and actually inappropriate. The Landis and Koch paper was specifically about the kappa statistic, not about the SNR statistic for reliability. It was not set in the context of health care quality measures, and there is no reason to presume that its rule of thumb guidance has any value in a context outside of kappa and agreement between or among raters. It's not necessary and helpful here, and actually detracts from, rather than enhances, the presentation.
11/15/21	Sharon L. Dunn, PT, PhD, American Physical Therapy Association (APTA)	Medical associations and societies	MIPS THA/TKA PRO-PM	On behalf of our 100,000 member physical therapists, physical therapist assistants, and students of physical therapy, the American Physical Therapy Association appreciates the opportunity to submit comments in response to the Centers for Medicare & Medicaid Services' and Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation's Development of a 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). APTA is dedicated to building a community that advances the physical therapy profession to improve the health of society. As experts in rehabilitation, prehabilitation, and habilitation, physical therapists play a unique role in society in prevention, wellness, fitness, health promotion, and management of disease, injury, and disability for individuals across the life span, helping individuals improve overall health and prevent the need for avoidable health care services. Physical therapists' roles include examination, diagnosis, management, education, direct intervention, research, advocacy, and collaborative consultation. These roles are essential to the profession's vision of transforming society by optimizing movement to improve the human experience. We appreciate the opportunity that Yale/CORE has given our member Craig Miller in serving as a Technical Expert Panel member for the development and reevaluation of orthopedic measures, including the development of the Clinician-Level and Clinician Group-Level THA/TKA PRO-PM. We believe APTA's participation in the TEP through Craig continues to be meaningful. We are grateful that Yale/CORE accepts input from all TEP members, and that Yale/CORE was amenable to reaching out to APTA regarding outcomes collection. APTA appreciates the well drafted Methodology Report, as well as the supportive document, which summarized the TEP meetings. We commend the various members of this panel for their time and expertise. We support the methodology and recommendations outlined within this document. In this comment opportunity, CMS encourages stakeholders to submit general or specific comments on the measure. APTA provides responses to each of CMS' questions below:

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/15/21	Continued from the previous page	Continued from the previous page	Continued from the previous page	• Is it preferable to report the measure at the clinician level or clinician-group level, or both, and why? It is preferable to report the measure at both the clinician and clinician-group levels. Significant change takes place when individual clinician practices combine into larger groups, so it would be best to report on both levels. There is a possibility that, as these trends change, reporting on both levels would not be as important. However, at this time, given the significant changes in practice for surgeons who transition from practicing in individual practice to joining larger groups, it is better to report the measure at both levels. • What are the preferences for approaches to transmitting patient-reported outcome measure data for this measure to CMS, such as via a clinical registry, third-party vendor, or other mechanism? APTA believes that transmitting PROM data for this measure to CMS would be best via some sort of clinical registry, as it would be both possible and cost-effective. Using a registry would assist the TEP in making data collection reliable and viable, which was a concern of the panel. • What is the optimal minimum response rate for measure calculation and/or future reporting? APTA supports the recommendation to expand the outcome data collection period to 300-425 days. This timeframe more accurately reflects postsurgical scheduling around the one-year milestone, and the increased window will not likely have an impact upon the overall outcome scores. APTA also believes that CMS/CORE should consider the wide variation in volume of rural versus urban versus suburban settings, as well as in individual clinician versus clinician group settings. • Is there any additional input on the measure or future implementation planning? APTA would be interested to know if there was any correlation to PROM and setting and ancillary service utilization. We advocate for using data and both patient-reported and performance-based outcomes measures, and we believe there is no one perfect or preferred measure. That being said, how do the various measures correlate for broader implications for improvements in access, appropriate utilization and cost, and overall functional and patient-reported outcomes? Additionally, APTA believes it is important that data is reported and presented in meaningful ways, both to enable beneficiaries to

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/15/21	Continued from the previous page	Continued from the previous page	Continued from the previous page	make educated decisions about their care, and to enable clinicians to reflect and improve on the quality of care that they provide. APTA agrees that individuals undergoing THA and TKA are ideal candidates for assessing PROs. We appreciate the in-depth review of the literature pertaining to the PROM for this population. Similar reviews were performed in the development of clinical practice guidelines for physical therapists caring for these patient populations. We support the recommendations to include the PRO instruments (VR-12, PROMIS-Mental Health, The HOOS JR, KOOS JR, Oswestry Index (back pain), as well as the questions pertaining to pain in the non-operative leg). We agree with using the clinical benefit thresholds identified by Lyman et al (22-point improvement with the HOOS JR and 20-point improvement for the KOOS JR). Although the TEP discussed the inconsistent findings regarding health literacy between years one and two with the TJR pilot, we do not believe the effects of this important domain on PRO has been adequately addressed. Research pertaining to health literacy levels, patient engagement, system-level literacy, and PRO scores is limited. We propose that the TEP make formal recommendations to further investigate the impact these disparities may have on outcomes for individuals undergoing total joint replacement. Additionally, APTA agrees with several considerations addressed by the TEP, including that the TEP should: • Expand the measure cohort to include outpatient and ambulatory surgical center settings given the expanding volume of patients in those environments for THA/TKA. APTA supports Yale/CORE's plans to test the measure using THA/TKA procedures that occur in the outpatient and ambulatory surgical center settings in the future. • Include considerations for addressing response bias. • Support the risk-adjustment model and consider social risk factors. While we were surprised that some of the social risk factors didn't seem to impact outcome, APTA appreciates the TEP's consideration for social risk factors. • Focus on developing an efficient and cost-effective way for data submission, especially considering variations in clinician and/or clinician group volumes.
11/17/21	Tilithia McBride, Vice President, Quality, Federation of American Hospitals (FAH)	Medical associations and societies	MIPS THA/TKA PRO-PM	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, what level of data collection burden for an individual PRO-PM is acceptable for a clinician, hospital, or other healthcare provider, and the degree to which duplicate data collection and reporting burden can be reduced with multiple

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/17/21	Continued from the previous page	Continued from the previous page	Continued from the previous page	groups (i.e., clinicians, practices, hospitals) implementing the same measure. Specifically, 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? Furthermore, if this measure is implemented in the Merit-based Incentive Payment System (MIPS), clinicians, practices and hospitals would be collecting and reporting on the same data but it does not appear that CMS sought to identify ways to mitigate the potential for duplicate data collection and reporting nor do we see any discussion on whether the duplication of effort could further increase the number of surveys or information that patients must provide. The FAH believes that these questions should have been addressed during the development of this PRO-PM and urges CMS to take the time to develop solutions to these questions prior to implementation of this measure in MIPS. In addition, 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

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/17/21	Continued from the previous page	Continued from the previous page	Continued from the previous page	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 appreciates the modifications that CMS made to the attribution approach to ensure that only those clinicians for whom it is appropriate to assign the procedure are included in the measure. We believe that it will further ensure the validity of the performance scores. The FAH urges CMS to address the critical issues around the feasibility of data collection and reduction of reporting burden prior to moving forward with this measure.
11/17/21	Richard Iorio, MD, President, American Association of Hip and Knee Surgeons (AAHKS); Michael J. Zarski, JD, Executive Director, AAHKS; Daniel K. Guy, MD, FAAOS, President, AAOS	Medical associations and societies	MIPS THA/TKA PRO-PM	The American Association of Hip and Knee Surgeons (AAHKS) appreciates the opportunity to submit comments to the Centers for Medicare & Medicaid Services (CMS) on the development of a clinician-level and clinician group-level THA and TKA PRO-PM. CMS seeks stakeholder feedback on the Project which has a goal to assess the quality of care provided to Medicare beneficiaries by clinicians and clinician groups performing elective primary THA and/or TKA procedures using patient-reported outcome (PRO) data. Overall, AAHKS supports the re-specification of the Hospital THA/TKA PRO-PM as a clinician-level and clinician group-level measure for the Quality Payment Program but believes that it is premature to introduce this measure at this time. PROs are the best available means for a patient-centered measurement of functional status improvement, the ultimate objective of arthroplasty; however, AAHKS has the following concerns. I. Whether it is preferable to report measures at the clinician-level or clinician group-level or both, and why? AAHKS supports reporting at both the clinician-level and clinician group-level. The opportunity to report as either a clinician or a clinician-group maximizes the opportunity for participation in the measure. However, the creation of sub-group reporting as a goal in 2026, as stated in the 2022 final PFS, will require new processes not currently in place. II. Preferences for approaches to transmitting patient-reported outcome measure (PROM) data for this measure to CMS (e.g., via a clinical registry, third party vendor)

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/17/21	Continued from the previous page	Continued from the previous page	Continued from previous page	AAHKS supports use of clinical registry submissions as a method for transmitting PROMs and believes that this is an opportunity to further encourage provider utilization of QCDRs, such as the American Joint Replacement Registry (AJRR) which is the registry primarily used for physician-level reporting for CJR, BPCI-A, and MIPS. PROM collection itself can be costly and time consuming and participation in data submission through a third-party vendor will result in additional costs to the clinician/clinician groups whose hospitals have chosen to not participate in a clinical registry. This presents a barrier to participation for those without infrastructure or hospital support for PROM collection and transmission. The same barrier exists for patients who are HOPD THA/TKA and those whose care is completed at an ASC. III. Optimal minimum response rate for measure calculation and/or future reporting The target for minimum response rate has no basis at this time. The CJR experience is not appropriate for determining collection rate targets given that there was significant incentive provided in terms of collection targets being met. Studies evaluating PROMs have demonstrated significant variation in response rate based on numerous demographic factors. Ideally data collection during voluntary collection with incentivization for high collection rates would precede final determination. IV. Any additional input on the measure for future implementation planning AAHKS supports clinician-level and clinician group-level PRO-PM for QPP, but believes that it is premature to introduce the measure at this time without having an appropriate benchmark for response rate. AAHKS recommends the initial implementation of the clinician-level and clinician group-level PRO-PM as a voluntary quality point rewarded structural measure to assess the viability of collection of PROMs by clinician and clinician groups which would then be used as a means of establishing the expected response rate. Stratified process targets starting at low levels would provide a minimum basis to establish the expected response rate and could be raised to find any potential ceiling. PROM collection and reporting is valuable and appropriate, but it is nevertheless an administrative burden on multiple parties that make a PROM capture rate of 60-70% a challenge. In particular, AAHKS is concerned for individual practitioners and small groups who will need additional time to develop policies and procedures to collect and report these data. Therefore, AAHKS recommends delaying inclusion of this measure in the QPP for 2 years to monitor and assess its use at the hospital level before expanding to the clinician level.

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/17/21	Continued from the previous page	Continued from the previous page	Continued from previous page	AAHKS welcomes the risk model developed for this measure but believes that ongoing evaluation of the risk factors in the model is required to ensure that the risk factors identified from CJR patients are appropriate when applied to the entire Medicare population. AAHKS has long favored the inclusion in risk adjustment of sociodemographic factors of patients. AAHKS recognizes that the statistics on the sample population of CJR patients did not demonstrate an influence on the PRO-PM, however, to mitigate even the perception of sociodemographic risk, AAHKS recommends stratification by proportions of dual-eligibility similar to what is now used by the CMS Readmission Reduction Program. If and when CMS proposes to formally add this PRO-PM to QPP, AAHKS supports an incentive for the collection PROMs. The development of a new E/M code to support collection of PROMs would support clinicians and clinician groups to develop infrastructure for data collection and support participation in registries for data submission. Once a target response rate has been established, an incentive program providing additional Quality Points for high performers will encourage C/CGs to improve systems for collection and submission making the measure more meaningful as response rates improve.
11/22/21	Koryn Rubin, Assistant Director of Federal Affairs, American Medical Association (AMA)	Medical associations and societies	MIPS THA/TKA PRO-PM	The American Medical Association (AMA) appreciates the opportunity to comment on the 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). The AMA supports the additional refinement made to the attribution approach. Specifically, other measures have experienced a similar issue of assigning patients to clinicians not primarily responsible for the procedure or episode such as physician assistants. We appreciate that this inaccurate attribution was identified and steps were taken to address the issue. We urge CMS to consider adopting this approach for other quality and cost measures in MIPS. 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.

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/22/21	Continued from previous page	Continued from previous page	Continued from previous page	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 the CMS adequately assessed the feasibility and potential data collection burden to the clinician, practice, and patient. Specifically, 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 published report does not 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 implementing this measure in MIPS. 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 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

Date Posted/ Received	Name, Credentials, and Organization of Commenter	Type of Organization	Measure Set or Measure	Text of Comments
11/22/21	Continued from previous page	Continued from previous page	Continued from previous page	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 CMS must provide additional information on these concerns particularly on how the feasibility of data collection and how the duplication of effort by hospitals, clinicians, and practices will be reduced prior to implementation of this measure in MIPS.

REFERENCES

1. Lyman, S., et al., *What Are the Minimal and Substantial Improvements in the HOOS and KOOS and JR Versions After Total Joint Replacement?* Clin Orthop Relat Res, 2018. **476**(12): p. 2432-2441.

2. Centers for Medicare and Medicaid Services. *Meaningful Measures Hub* 2021; Available from: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/MMF/General-info-Sub-Page.