NATIONAL QUALITY FORUM
Driving measurable health improvements together

# Memo

**September 10, 2020**

**To**: Patient Experience and Function Standing Committee

**From**: NQF staff

**Re**: Post-comment web meeting to discuss public comments received and NQF member expression of support

## Purpose of the Call

The Patient Experience and Function Standing Committee will meet via web meeting on September 17, 2020 from 11:00 am – 1:00 pm ET. The purpose of this call is to:

- Review and discuss comments received during the post-evaluation public and member comment period;
- Provide input on proposed responses to the post-evaluation comments;
- Review and discuss NQF members' expression of support of the measures under consideration; and
- Determine whether reconsideration of any measures or other courses of action are warranted.

## Standing Committee Actions

1. Review this briefing memo and draft report.
2. Review and consider the full text of all comments received and the proposed responses to the post-evaluation comments (see comment table).
3. Review the NQF members' expressions of support of the submitted measures.
4. Be prepared to provide feedback and input on proposed post-evaluation comment responses.

## Conference Call Information

Please use the following information to access the conference call line and webinar:

**Dial-in #:** 1-800-768-2983
**Access code:** 4364232
**Web link:** https://core.callinfo.com/callme/?ap=8007682983&ac=4364232&role=p&mode=ad

## Background

Over the past decade, there have been increasing efforts to change the healthcare paradigm from one that identifies persons as passive recipients of care to one that empowers individuals to participate actively in their care.[1-3] Healthcare treatments can be tailored to individual patients in terms of patient preferences and individual clinical factors when the patient voice is captured as part of routine care. Capturing patient experience and evaluating patient function are two important components of patient-centered measurement.[4]

The Centers for Medicare & Medicaid Services (CMS) Meaningful Measures Initiative includes the identification of measures that capture patients' experiences with clinicians and providers—one of 19 measurement areas for focusing our healthcare quality improvement efforts as a country.[5] This falls under the measurement priority associated with strengthening person and family engagement as partners in their care. Ensuring that each person and family is engaged within a care partnership is critical to achieving better patient outcomes.[6]

Patient Experience and Function (PEF) is a National Quality Forum (NQF) measure topic area encompassing patient functional status, satisfaction, and experience of care, as well as issues related to care coordination. Central to the concepts associated with patient experience with their overall care is the patient's health-related quality of life and many factors that influence it, including communication, care coordination, transitions of care, and use of health information technology.[7-9]

The 25-member Patient Experience and Function Standing Committee has been charged with overseeing the NQF patient experience and function measure portfolio. The Committee evaluates both newly submitted and previously endorsed measures against NQF's measure evaluation criteria, identifies gaps in the measurement portfolio, provides feedback on how the portfolio should evolve, and serves on any ad hoc or expedited projects in its designated topic areas.

During the June 23 and June 24, 2020 web meetings, the Patient Experience and Function Standing Committee evaluated four measures undergoing maintenance review. The Standing Committee recommended one new measure for endorsement and recommended three measures for continued endorsement.

The new measure recommended by the Committee for endorsement is:

- 3559 Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA) (Centers for Medicare & Medicaid Services (CMS))

The measures the measures recommended for continued endorsement are:

- 2614 CoreQ: Short Stay Discharge Measure (AHCA/NCAL)
- 2615 CoreQ: Long-Stay Resident Measure (American Health Care Association)
- 2616 CoreQ: Long-Stay Family Measure (AHCA/NCAL)

## Comments Received

NQF solicits comments on measures undergoing review in various ways and at various times throughout the evaluation process. First, NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). Second, NQF solicits member and public comments during a 16-week comment period via an online tool on the project webpage.

### Pre-evaluation Comments

NQF solicits comments prior to the evaluation of the measures via an online tool on the project webpage. For this evaluation cycle, the pre-evaluation comment period was open from May 1, 2020 to June 12, 2020 for the measures under review. NQF received one comment on *NQF 3559: Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA)* related to multi-step inclusion of social risk factors in risk adjustment model, feasibility i.e. cost of implementation, potential survey fatigue to patients, and process for and administrative burden of data collection. This pre-evaluation comment will be provided to the Committee during the post-comment web meeting.

## Comments and Their Disposition

### *Measure-Specific Comment*

**3559: Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA)**

The Federation of American Hospitals (FAH) appreciates the opportunity to comment on measure #3559, Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA), prior to the Standing Committee's evaluation. The 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. For instance, the degree to which multiple PRO-PMs could lead to survey fatigue for patients, the potential impact additional PRO-PMs may have on the reporting of well-established measures such as HCAHPs, and what level of data collection burden for an individual PRO-PM is acceptable for a hospital or other healthcare provider.

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 hospitals such as whether it required coordination with orthopedic practices 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 hospital 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? The FAH believes that these questions should have been addressed during the development of this PRO-PM and this detail should have been provided within the measure submission rather than the generalized statements that we see in the responses under the feasibility criterion.

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 the measure developer 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 hospitals 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.

**Measure Steward/Developer Response:**

The Federation of American Hospitals submitted a public comment on June 12, 2020 for Measure 3559, stating support for PRO-PM development and implementation but noting concerns about data collection and hospital burden, survey fatigue for patients, the proposed measure's impact on response to other established PRO-PMs such as HCAHPS, and the inclusion of social risk variables. We'd like to thank the Federation of American Hospitals for their comment and have provided responses to their concerns below.

The Centers for Medicare and Medicaid Services (CMS) is in the process of developing a strategic plan for implementing this measure into public reporting, informed by stakeholder input and with careful consideration of hospital burden. While PRO-PMs require providers to integrate data collection into clinical workflows, this integration provides opportunity for PROs to inform clinical decision making and benefit patients by engaging them in discussions about outcomes. The manner in which data collection is integrated into clinical workflows will be at the discretion of the hospital, which allows them freedom and flexibility to choose an approach that meets the needs of their facility.

We do not expect this PRO-PM to contribute to survey fatigue or to negatively impact other PRO-PMs. The Patient-Reported Outcome Measure (PROM) instruments that are used to calculate pre- and postoperative scores for this THA/TKA PRO-PM were carefully selected, with extensive stakeholder input, to be low burden for patients. Each patient undergoing a primary elective THA/TKA will be asked to fill out one joint-specific PROM (6 questions on the HOOS, JR for hip replacement patients and 7 questions on the KOOS, JR for knee replacement patients) and one global health PROM (10 to 14 questions). In addition, it is unlikely that patients would be put in a position to prioritize the THA/TKA PRO-PM under consideration over other established PRO-PMs, such as the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS), as the assessment timelines for these two measures are completely different, with no overlap. The PRO data for the PRO-PM under consideration are collected 90 to 0 days before undergoing an elective, primary THA/TKA and 270 to 365 days after the THA/TKA, whereas the HCAHPS survey is administered to a random sample of adults within 2 days to 6 weeks after hospital discharge. Additionally, while sample sizes and patient caseloads vary by hospitals, we anticipate based on HCAHPS minimum caseload (for example, 100 completed surveys in a typical annual Hospital Value Based Purchasing program performance period) that most patients receiving the PROM surveys for this THA/TKA PRO-PM would not have also completed the HCAHPS survey for their hospital stay. Nationally, THA/TKA procedures comprise a small portion of hospital inpatient discharges, so we anticipate that relatively few patients would have completed both survey instruments covering their hospital stays.

The risk model for Measure 3559 was developed with clinical risk factors identified through a systematic literature review and orthopedic input. In consultation with the Technical Expert Panel (TEP), clinical consultants, and through detailed public comments, including from specialty societies, we focused on candidate risk-adjustment variables of interest that were clinically relevant and had an evidence-based relationship with clinical outcomes following elective primary THA or TKA. This included health literacy, identified by clinical experts and stakeholders as an important predictor and associated with a range of social risk factors. While health literacy is associated with other aspects of social risk, our TEP and measure development team felt that health literacy held particular relevance for a measure based upon PRO data. Thus, health literacy was evaluated during the initial risk model development along with clinical candidate risk variables. An additional assessment of the impact of social risk as captured by dual

eligibility, the Agency for Health Research and Quality (AHRQ) socioeconomic status (SES) Index (socioeconomic status), and non-white race, was performed after the risk model was finalized. The addition of each of these three social risk variables provided no statistically significant change to our model performance, variable coefficients, or the model outcome. These variables were not included in the present model, but we believe future assessment in reevaluation will be important. These social risk variables were, however, statistically significantly associated with response to PRO surveys—whether patient-reported outcomes were obtained for patients undergoing primary elective THA/TKA—and so were included in the calculation of stabilized inverse probability weights used to account for potential response-bias.

**Proposed Committee Response:**

Thank you for your comments. The Committee will review these comments during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and the developer's response and be prepared to discuss any recommendations for the developer to consider.

## Post-evaluation Comments

The draft report was posted on the project webpage for public and NQF member comment on August 5, 2020 for 30 calendar days. During this commenting period, NQF received five comments from three member organizations:

| Member Council | # of Member Organizations Who Commented |
|---|---|
| Supplier/Industry | 1 |
| Consumer | 1 |
| Health Profession | 2 |

We have included all post-evaluation comments that we received in the comment table (Excel spreadsheet) posted to the Committee SharePoint site. This comment table contains the commenter's name, comment, associated measure, topic (if applicable), and—for the post-evaluation comments—draft responses (including measure steward/developer responses) for the Committee's consideration. Please review this table in advance of the meeting and consider the individual comments received and the proposed responses to each.

## Comments and Their Disposition

### *Themed Comments*

One theme was identified in the post-evaluation comment period:

1. Consensus Development Process

## Theme 1 – Consensus Development Process

The American Medical Association and the Federation of American Hospitals noted concerns with the Consensus Development Process (CDP). Both stakeholders acknowledged that NQF 2614 – 2616 originate from the same developer and use similar specifications and testing approaches, but they

disagreed with the Standing Committee's motion to carry evaluation criteria ratings from NQF 2614 to 2615 and subsequently, NQF 2615 to 2616 with no discussion. They disagreed with the notion that the evidence, reliability, validity, feasibility, use and usability are similar enough across NQF 2614 – 2616 to carry votes forward with no discussion, and they stated that the evidence, sampling strategies, testing results, target populations, different survey tools, feasibility, and usability differ enough to prompt discussion prior to a vote. The AMA and FAH requested clarification on these voting actions that were taken by the Standing Committee for NQF 2615 and 2616.

An additional process concern was raised, specifically by the Federation of American Hospitals, about an omission of a pre-evaluation comment in the Spring 2020 Patient Experience and Function Draft Report. Additionally, a request to make the pre-evaluation comment available to the Standing Committee was included.

**NQF Response**

NQF wishes to thank the commenters for the procedural question related to the carrying of the vote for measures 2614-2616. When considering measures with comparable submissions, NQF Committees may elect to carry the vote on a given criteria provided that there is unanimous support for doing so; a single dissention is sufficient to move to a full vote. This is not a departure from NQF's consensus develop process (CDP). In fact, due to the concerns also identified by FAH, it requires a much higher degree of consensus than other CDP procedures. It is also not uncommon. The same measures were passed using the option to carry the vote for certain criteria during their previous evaluation in 2016, as were several other measures since.

**Measure Steward/Developer Response:**

No response is required.

**Proposed Committee Response:**

Thank you for your comments. The Committee will review these comments during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and the NQF staff response and be prepared to discuss any recommendations for the developer to consider.

### *Measure-Specific Comments*

#### 2614 Core Q: Short Stay Discharge Measure

The American Geriatrics Society (AGS) wishes to provide comment on Measure 2614. The exclusion of those who are readmitted to acute care, transferred to another skilled nursing facility (SNF) or long-term acute care (LTAC) facility, or remain in an SNF longer than 100 days removes subsets of patients who are more likely to have lower satisfaction with their short term SNF stay. We believe that these exclusions make the current measure less meaningful. We have additional reservations about the chosen cut-point for the score as 3 or above. A response of 2 is designated as 'average' in the Core Q questionnaire. We believe that this is not necessarily a negative response and could still indicate satisfaction with care.

**Measure Steward/Developer Response:**

Thank you for taking the time to review and comment on our measure. AHCA/NCAL greatly values feedback from patients, family members, and the public. We believe it to be paramount to measure the satisfaction of short and long-stay residents as well as family members (or designated parties), which is why we developed and tested measures NQF# 2615 ( CoreQ: Long-stay resident measure) and NQF# 2616 (CoreQ: Long-stay family measure), to assess the satisfaction of long-stay residents and their family members/designated parties. The scoring of

the CoreQ was tested using different cut-points. Facilities were very stable in their position relative to others when different cut-points were used. However, with a cut-point of two, many more facilities score at the highest levels. As such, if they subsequently improve their satisfaction, this is not well reflected in the change in CoreQ score. For this purpose, a cut-point of 3 works best. Also, "average" may not necessarily be a "bad" score - but with input from nursing home representatives during the testing of the CoreQ scoring, it was considered important to move the industry beyond average.

**Proposed Committee Response:**
Thank you for your comments on our review of NQF 2614. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**
The Committee should review the comments and the developer's response and be prepared to discuss any recommendations for the developer to consider.

### 2615 Core Q: Long-Stay Resident Measure

### 2616 Core Q: Long-Stay Family Measure

The American Geriatrics Society (AGS) wishes to provide comment on Measure 2615 and 2616. We have reservations about the chosen cut-point for the score as 3 or above. A response of 2 is designated as 'average' in the Core Q questionnaire. We believe that this is not necessarily a negative response and could still indicate satisfaction with care.

**Measure Steward/Developer Response:**
Thank you for taking the time to review and comment on our measure. AHCA/NCAL greatly values feedback from patients, family members, and the public. The scoring of the CoreQ was tested using different cut-points. Facilities were very stable in their position relative to others when different cut-points were used. However, with a cut-point of 2 many more facilities score at the highest levels. As such, if they subsequently improve their satisfaction this is not well reflected in the change in CoreQ score. For this purpose, a cut-point of 3 works best. Also, "average" may not necessarily be a "bad" score - but with input from nursing home representatives during the testing of the CoreQ scoring, it was considered important to move the industry beyond average.

**Proposed Committee Response:**
Thank you for your comments on our review of NQF 2615-16. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**
The Committee should review the comments and the developer's response and be prepared to discuss any recommendations for the developer to consider.

### 2614 Core Q: Short Stay Discharge Measure

### 2615 Core Q: Long-Stay Resident Measure

### 2616 Core Q: Long-Stay Family Measure

I am commenting as a Patient/Public individual interested and engaged in quality measurement and improvement. Thank you for offering the public opportunity to comment on this PEF NQF set of

measures. I find the 3 QC measures appropriate for use, though would like to add that I have heard on a number of occasions that patient experience does not = patient satisfaction. Patient satisfaction is the outcome measured in these three measures. Are these designed to be used in place of CAHPS as they are more specific to stays in LTC/SNFs etc.? Family member identified for measurement query might vary as well. One family member might be pleased another not so much, just food for thought.

**Measure Steward/Developer Response:**

Thank you for taking the time to review and comment on our measure. AHCA/NCAL greatly values feedback from patients, family members, and the public. We agree that experience and satisfaction are two different, albeit related, constructs. NQF #2614: CoreQ: Short Stay Discharge Satisfaction measure, was specified and tested to be a patient satisfaction measure, only. NQF# 2615: CoreQ: Long-Stay Resident Satisfaction measure, was specified and tested to be a resident satisfaction measure, only. NQF# 2616: CoreQ: Long-Stay family measure, was specified and tested to be a family/designated party satisfaction measure, only.

**Proposed Committee Response:**

Thank you for your comments on our review of NQF 2614 - 2616. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and the developer's response and be prepared to discuss any recommendations for the developer to consider.

### 3559: Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA)

The American Medical Association (AMA) appreciates the opportunity to comment and vote on NQF #3559: Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA). The AMA supports the assessment of patient-reported outcomes but believes that the burden of data collection both to the hospital and the patient must be adequately addressed and the continued multi-step approach to risk adjustment must be reconsidered. On review of the draft report, we were unable to find sufficient discussion of the validity and usability of this measure and ask that the Committee reconsider the proposed endorsement of this measure in light of our concerns outlined below.

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 and that these data are 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 developer (CMS) adequately assessed the feasibility and potential data collection burden both to the hospital and patient. Specifically, the responses to the questions on feasibility do not discuss how the testing sites coordinated data collection across settings or whether the responsibility of the additional items was placed on the hospital. This question is particularly important since the specifications require hospitals to collect data for one measure from 90 days pre-operatively to up to one-year post-operative. Perhaps more importantly, we would have liked 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. For example, if these data were collected on the morning of the surgery, could stress and anxiety have impacted responses or 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? 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 and hospitals as more and more patient-reported outcome performance measures 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,[1] 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 not recommend the measure for endorsement until they are adequately addressed.

[1] National Quality Forum. *Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors*. Final report. July 18, 2017. Available at: http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85635. Last accessed December 18, 2018.

**Measure Steward/Developer Response:**

The American Medical Association submitted a public comment on August 25, 2020 for Measure 3559, stating support for PRO-PM development but noting concerns about data collection and hospital burden, the inclusion of social risk variables, intended use of the measure, and sufficient NQF discussion regarding the validity of the measure. We'd like to thank the American Medical Association for their comment and have provided responses to their concerns below.

The issue of burden has been carefully considered during measure development and will be carefully considered for measure implementation. The Centers for Medicare and Medicaid Services (CMS) is in the process of developing a strategic plan for implementing this measure into public reporting, informed by stakeholder input and with careful consideration of hospital burden. While patient-reported outcome performance measures (PRO-PMs) require providers to integrate data collection into clinical workflows, this integration provides opportunity for PROs to inform clinical decision making and benefit patients by engaging them in discussions about outcomes. The manner in which data collection is integrated into clinical workflows will be at the discretion of the hospital, which allows them freedom and flexibility to choose an approach that meets the needs of their facility. CMS's Comprehensive Care for Joint Replacement (CJR) program, the source of the data for measure development and testing, served as proof of concept for patient-reported outcome data collection and submission.

We do not expect this PRO-PM to contribute to survey fatigue or to negatively impact the collection of data for other PRO-PMs. The Patient-Reported Outcome Measure (PROM) instruments that are used to calculate pre- and postoperative scores for this THA/TKA PRO-PM were carefully selected, with extensive stakeholder input, to be low burden for patients. Each patient undergoing a primary elective THA/TKA will be asked to fill out one joint-specific PROM (6 questions on the HOOS, JR for hip replacement patients and 7 questions on the KOOS, JR for knee replacement patients) and one global health PROM (10 to 14 questions). In addition, it is unlikely that patients would be less likely to complete other PROM surveys such as the Hospital

Consumer Assessment of Healthcare Providers and Systems (HCAHPS), as the assessment timelines for these two measures are completely different, with no overlap. As noted, the PRO data for the PRO-PM under consideration are collected 90 to 0 days before undergoing an elective, primary THA/TKA and 270 to 365 days after the THA/TKA, whereas the HCAHPS survey is administered to a random sample of adults within 2 days to 6 weeks after hospital discharge. Additionally, while sample sizes and patient caseloads vary by hospitals, we anticipate based on HCAHPS minimum caseload (for example, 100 completed surveys in a typical annual Hospital Value Based Purchasing program performance period) that most patients receiving the PROM surveys for this THA/TKA PRO-PM would not have also completed the HCAHPS survey for their hospital stay. Nationally, THA/TKA procedures comprise a small portion of hospital inpatient discharges, so we anticipate that relatively few patients would have completed both survey instruments covering their hospital stays.

And while the addition of PRO-PMs over time will require continued consideration of patient and hospital burden, the value of patient-reported outcomes to patient-centered care, improved patient care and improved hospital quality is widely supported. This measure was developed with extensive input from patients, who indicated strong support for a patient-reported outcomes-based performance measure following elective primary THA and TKA. They stated that patients expected a significant amount of improvement in pain levels and functional status, which helped define improvement thresholds for the measure outcome. They reviewed improvement threshold considerations, the preoperative and postoperative data collection periods, risk adjustments models, and testing results, and found the measure to be valuable. They provided feedback on the presentation of the measure outcome for patient use.

The risk model for Measure 3559 was developed with clinical risk factors identified through a systematic literature review and orthopedic input. In consultation with the Technical Expert Panel (TEP), clinical consultants, and through detailed public comments, including from specialty societies, we focused on candidate risk-adjustment variables of interest that were clinically relevant and had an evidence-based relationship with clinical outcomes following elective primary THA or TKA. This included health literacy, identified by clinical experts and stakeholders as an important predictor and associated with a range of social risk factors. While health literacy is associated with other aspects of social risk, our TEP and measure development team felt that health literacy held particular relevance for a measure based upon PRO data. Thus, health literacy was evaluated during the initial risk model development along with clinical candidate risk variables. An additional assessment of the impact of social risk as captured by dual eligibility, the Agency for Health Research and Quality (AHRQ) socioeconomic status (SES) Index (socioeconomic status), and non-white race, was performed after the risk model was finalized. The addition of each of these three social risk variables provided no statistically significant change to our model performance, variable coefficients, or the model outcome. These variables were not included in the present model, but we believe future assessment in reevaluation will be important. These social risk variables were, however, statistically significantly associated with response to PRO surveys—whether patient-reported outcomes were obtained for patients undergoing primary elective THA/TKA—and so were included in the calculation of stabilized inverse probability weights used to account for potential response-bias. We will continue to assess the impact of social risk factors on measure over time.

Finally, extensive discussion regarding the validity of measure was conducted by the Scientific Methods Panel, which subsequently voted in support of the measure on both reliability and validity.

**Proposed Committee Response:**

Thank you for your comments on our review of NQF 3559. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and the developer's response and be prepared to discuss any recommendations for the developer to consider.

The American Geriatrics Society (AGS) wishes to provide comment on Measure 3559. While this seems like a potentially useful PROM it's unclear how it would be implemented. There should be some consideration given to the amount of time hospital staff would devote to collecting this measure as well as the costs involved, both of which will be borne directly by the hospitals. We suggest that this measure could be replaced by data which hospitals may already be collecting (such as ADLs, IADLs, or pain scores) and therefore do not pose as much of a burden on reporting hospitals.

**Measure Steward/Developer Response:**

The American Geriatrics Society submitted a public comment on September 3, 2020 for Measure 3559, stating support for the potential usefulness of the measure but expressing concern about implementation and hospital burden. We'd like to thank the American Geriatrics Society for their comment and have provided responses to their concerns below.

The Centers for Medicare and Medicaid Services (CMS) is in the process of developing a strategic plan for implementing this measure into public reporting, informed by stakeholder input and with careful consideration of hospital burden. While patient-reported outcomes performance measures (PRO-PMs) require providers to integrate data collection into clinical workflows, this integration provides opportunity for patient reported outcomes (PROs) to inform clinical decision making and benefit patients by engaging them in discussions about potential outcomes. The manner in which data collection is integrated into clinical workflows will be at the discretion of the hospital, which allows them freedom and flexibility to choose an approach that meets the needs of their facility.

This performance measure, built on established and validated PROM survey instruments, will standardize across hospitals the data collected, the assessment periods for data collection, and the risk adjustment used for evaluation of hospital quality following elective primary THA and TKA procedures. While most hospitals do collect functional and pain information from orthopedic patients (such as ADLs, IADLS, and pain scores), assessment of hospital-level performance for public reporting requires a standard approach across all hospitals. The patient reported outcome measure (PROM) instruments that are used to calculate pre- and postoperative scores for this THA/TKA PRO-PM were carefully selected, with extensive stakeholder input from clinicians, to be low burden and joint specific. The clinicians felt, and data demonstrated, that joint-specific functional status tools such as the HOOS, JR and KOOS, JR are more relevant for clinical decision making and are more responsive than PROMs that are not as specific. In addition, this measure was developed with extensive input from patients, who indicated strong support for a PRO-PM following elective primary THA and TKA. They stated that patients expected a significant amount of improvement in pain levels and functional status, which helped define improvement thresholds for the measure outcome. They reviewed improvement threshold considerations, the preoperative and postoperative data collection periods, risk adjustments models, and testing results. Patients provided feedback on the presentation of the measure outcome for patient use and found the measure to be valuable.

**Proposed Committee Response:**

Thank you for your comments on our review of NQF 3559. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and be prepared to discuss any recommendations for the developer to consider.

As a patient/public partner engaged in quality improvement, my comments are as follows. I find it interesting that #3559 measure did not have an intended use identified pre NQF endorsement phase. It seems to me that having the use case open allows for ease for endorsement though also some might feel that the intended use of measures should be identified in development therefore decreasing opportunities for questions later. I did appreciate the longevity of time for follow up as in these types of total hip/knee arthroplasty surgeries, the 30/60/90-day outcomes might not reflect actual patient experience or function post-surgery that impacts their quality of life. Thank you for this.

**Measure Steward/Developer Response:**

An independent member of the public submitted a public comment for Measure 3559. They expressed support for the measure and the timeframe for post-operative data collection, noting that the longer follow up is beneficial for measuring patient experience and quality of life. They also expressed a concern about the intended use of the measure, which was not identified during the development of the measure. We would like to thank this individual for their comment and have provided a response to their concern below.

During NQF evaluation of new measures, fully identified “Usability and Use” is not required for NQF endorsement. The Centers for Medicare and Medicaid Services (CMS) is in the process of developing a strategic plan for implementing this measure into public reporting, informed by stakeholder input and with careful consideration of hospital burden. CMS’s Comprehensive Care for Joint Replacement (CJR) program, the source of the data for measure development and testing, served as proof of concept for patient-reported outcome data collection and submission.

This measure was developed with extensive stakeholder input and support from patients, orthopedic groups, and clinical experts. Patients indicated strong support for a patient-reported outcomes-based performance measure following elective primary THA and TKA. They stated that patients expected a significant amount of improvement in pain levels and functional status, which helped define improvement thresholds for the measure outcome. They reviewed improvement threshold considerations, risk adjustments models, and testing results, and found the measure to be valuable. They provided feedback on the presentation of the measure outcome for patient use.

**Proposed Committee Response:**

Thank you for your comments on our review of NQF 3559. The Committee will review this comment during its deliberations on the Post-Comment Call scheduled on September 17, 2020.

**Action Item:**

The Committee should review the comments and be prepared to discuss any recommendations for the developer to consider.

## NQF Member Expression of Support

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ("support" or "do not support") for each measure submitted for endorsement consideration to inform the Committee's recommendations. One NQF member provided their expression of support/non-support: See Appendix A.

## Appendix A: NQF Member Expression of Support Results

One NQF member provided their expressions of support/non-support. None of the measures under consideration received support from NQF members. Results for the measure that received expressions of support/non-support are provided below.

**3559: Hospital-Level, Risk-Standardized Patient-Reported Outcomes Following Elective Primary Total Hip and/or Total Knee Arthroplasty (THA/TKA)**

| Member Council | Support | Do Not Support | Total |
|---|---|---|---|
| Health Profession | 0 | 1 | 1 |

## References

[1] CMS. *Medicare Beneficiary Characteristics*. https://www.cms.gov/Research-Statistics-Data-andSystems/Statistics-Trends-and-Reports/Chronic-Conditions/Medicare_Beneficiary_Characteristics. Last accessed February 2020.

[2] United Health Foundation. America's Health Rankings website. https://www.americashealthrankings.org/explore/senior/measure/hospital_readmissions_sr/state/ALL. Last accessed February 2020.

[3] Institute of Medicine Roundtable on Evidence-Based Medicine; Yong PL, Saunders RS, Olsen LA, eds. *The Healthcare Imperative: Lowering Costs and Improving Outcomes: Workshop Series Summary*. Washington, DC: National Academies Press; 2010.

[4] Agency for Healthcare Research and Quality (AHRQ). Priorities of the national quality strategy website. https://www.ahrq.gov/research/findings/nhqrdr/nhqdr15/priorities.html. Last accessed February 2020.

[5] Centers for Medicare & Medicaid Services. Meaningful Measures Hub website. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-AssessmentInstruments/QualityInitiativesGenInfo/MMF/General-info-Sub-Page#Measurement_Areas. Last accessed February 2020.

[6] Frandsen BR, Joynt KE, Rebitzer JB, et al. Care fragmentation, quality, and costs among chronically ill patients. *AJMC*. 2015; 21(5):355-362.

[7] Schultz EM, Pineda N, Lonhart J, et al. A systematic review of the care coordination measurement landscape. *BMC Health Serv Res*. 2013;13:119.

[8] Rosenthal MB, Alidina S, Friedberg MW, et al. A difference-in-difference analysis of changes in quality, utilization and cost following the Colorado multi-payer patient-centered medical home pilot. *J of Gen Intern Med*. 2016;31(3):289-296.

[9] Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *New Engl J Med*. 2009;360(14):1418-1428.