



## Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

### Brief Measure Information

**NQF #:** 2653

**Corresponding Measures:**

**Measure Title:** Average change in functional status following total knee replacement surgery

**Measure Steward:** MN Community Measurement

**sp.02. Brief Description of Measure:** For patients age 18 and older undergoing total knee replacement surgery, the average change from pre-operative functional status to one year (nine to fifteen months) post-operative functional status using the Oxford Knee Score (OKS) patient reported outcome tool.

**1b.01. Developer Rationale:** Annually there are over 500,000 total knee replacement (TKR) procedures performed in the US. It is projected that by 2030 the volume of this procedure will increase to over 3.48 million per year due to the aging baby-boomers, increased obesity and indications for TKR that extend to both younger as well as older patients. From 2000 to 2006, the Medicare TKR rate overall in the United States increased 58%, from 5.5 to 8.7 per 1000 and TKR revisions currently represent 8.2% of all Medicare dollars spent. It is estimated that annual hospital charges for TKR will approach 40.8 billion dollars annually by 2015.

As Dartmouth Atlas data demonstrate there are significant differences in utilization of total knee replacement procedures for MN as compared to the national rates of utilization for Medicare enrollees. With the anticipated explosion in the number of procedures projected, associated total costs, lack of clinical guidelines/appropriateness criteria; MN's Measurement and Reporting Committee (MARC) and MN Board approved measure development focused on outcomes of functional status and quality of life following total knee replacement procedures to fill a gap in the current measure portfolio.

---

**sp.12. Numerator Statement:** There is not a traditional numerator for this measure; the measure is calculating the average change in functional status score from pre-operative to post-operative functional status score. The measure is NOT aiming for a numerator target value for a post-operative OKS score.

For example:

The average change in knee function was an increase of 15.9 points one year post-operatively on a 48 point scale.

**sp.14. Denominator Statement:** Adult patients age and older (no upper age limit) who undergo a primary or revision total knee replacement procedure during a calendar year performance period (e.g. dates of procedure occurring between 1/1/2013 and 12/31/2013) AND have a completed pre-operative and post-operative OKS patient reported outcome assessments.

**sp.16. Denominator Exclusions:** There are no denominator exclusions from the initial patient population for this measure.

---

**Measure Type:** Outcome: PRO-PM

**sp.28. Data Source:**

Paper Medical Records

Other

Instrument-Based Data

**sp.07. Level of Analysis:**

Clinician: Group/Practice

---

**IF Endorsement Maintenance – Original Endorsement Date:** 2015-07-07 04:30 PM

**Most Recent Endorsement Date:** 7/7/2015 4:30:28 PM

---

**IF this measure is included in a composite, NQF Composite#/title:**

**IF this measure is paired/grouped, NQF#/title:**

**sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?:**

## 1. Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria

**1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.**

[Response Begins]

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

**2021 Submission:**

Updated evidence information here.

**2018 Submission:**

Evidence from the previous submission here.

**1a.01. Provide a logic model.**

*Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.*

[Response Begins]

[Response Ends]

**1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.**

*Describe how and from whom input was obtained.*

[Response Begins]

[Response Ends]

**1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.**

[Response Begins]

[Response Ends]

**1b.01. Briefly explain the rationale for this measure.**

*Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.*

**[Response Begins]**

Annually there are over 500,000 total knee replacement (TKR) procedures performed in the US. It is projected that by 2030 the volume of this procedure will increase to over 3.48 million per year due to the aging baby-boomers, increased obesity and indications for TKR that extend to both younger as well as older patients. From 2000 to 2006, the Medicare TKR rate overall in the United States increased 58%, from 5.5 to 8.7 per 1000 and TKR revisions currently represent 8.2% of all Medicare dollars spent. It is estimated that annual hospital charges for TKR will approach 40.8 billion dollars annually by 2015.

As Dartmouth Atlas data demonstrate there are significant differences in utilization of total knee replacement procedures for MN as compared to the national rates of utilization for Medicare enrollees. With the anticipated explosion in the number of procedures projected, associated total costs, lack of clinical guidelines/appropriateness criteria; MN's Measurement and Reporting Committee (MARC) and MN Board approved measure development focused on outcomes of functional status and quality of life following total knee replacement procedures to fill a gap in the current measure portfolio.

**[Response Ends]**

**1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.**

*Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*

**[Response Begins]**

Two phases of pilot testing of measures was completed in July of 2014. Barriers to PRO tool administration were identified during pilot testing and an interim plan is in place to increase PRO administration rates. The functional status outcome measure is planned for public reporting mid-2016.

**Analysis of TKR Functional Status Outcomes \***

25 Medical Groups/ Practices; 7 orthopedic practices and 18 multi-specialty practices with a department or division of orthopedics (e.g. HealthPartners, Mayo Clinic). The 25 practices are further described by geographic location (5 metropolitan, 6 micropolitan, and 14 rural) and by annual volume of TKR procedures.

TKR volume < 100 8 practices 109 patients/ procedures; primary TKR

TKR volume 100 to 299 6 practices 203 patients/ procedures; primary TKR

TKR volume 300 to 600 3 practices 858 patients/ procedures; primary TKR

TKR volume > 600 5 practices 790 patients/ procedures; primary TKR

2,044 patients/ procedures. 1,960 primary TKR, 84 revision TKR

**Primary Total Knee Replacement**

Annual Volume TKR Average Change in OKS Score at 1 Year

Practices with < 100 14.4

Practices with 100 to 299 16.5

Practices with 300 to 600 17.2

Practices with > 600 17.3

Total 17.0 increase in points on a 48 point scale

\* Patients with completed pre-operative and post-operative OKS patient reported outcome tools

[Response Ends]

**1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.**

[Response Begins]

[Response Ends]

**1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.**

*Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.*

[Response Begins]

[Response Ends]

**1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.**

[Response Begins]

Racial Disparities in Total Knee Replacement Among Medicare Enrollees --- United States, 2000--2006  
An estimated 45% of U.S. adults might be at risk for developing symptomatic knee osteoarthritis during their lifetimes, with whites and blacks at equal risk for this common disabling condition. Total knee replacement (TKR) is an effective method of reducing pain and improving physical function among those with disabling knee osteoarthritis; however, whites have been more likely to undergo the procedure than blacks. As a result, a Healthy People 2010 objective\* calls for eliminating racial disparities in the rate of TKR among persons aged >65 years. To monitor progress toward achieving this objective, CDC analyzed national and state TKR rates for Medicare enrollees for the period 2000--2006, stratified by sex, age group, and black or white race. From 2000 to 2006, the TKR rate overall in the United States increased 58%, from 5.5 to 8.7 per 1,000 population, with similar increases among whites (61%) and blacks (56%). However, the TKR rate for blacks was 37% lower than the rate for whites in 2000 (3.6 versus 5.7 per 1,000 population) and 39% lower in 2006 (5.6 versus 9.2 per 1,000 population). Health-care providers and public health agencies might help reduce this disparity by widely distributing TKR information that is tailored to the education and literacy levels and culture of patients with symptomatic knee osteoarthritis. Health-care providers should conduct, as routine practice, thorough discussions regarding knee pain symptoms and loss of physical function with older patients of all races who might be candidates for TKR.  
Centers for Disease Control MMWR Morbidity and Mortality Weekly Report (MMWR) Feb 20 2009

[Response Ends]

## 2. Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.

**spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.**

[Response Begins]

[Response Ends]

**spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.**

**For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.**

*For example, specifications may have been updated based on suggestions from a previous NQF CDP review.*

[Response Begins]

[Response Ends]

**sp.01. Provide the measure title.**

*Measure titles should be concise yet convey who and what is being measured (see [What Good Looks Like](#)).*

[Response Begins]

Average change in functional status following total knee replacement surgery

[Response Ends]

**sp.02. Provide a brief description of the measure.**

*Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).*

[Response Begins]

For patients age 18 and older undergoing total knee replacement surgery, the average change from pre-operative functional status to one year (nine to fifteen months) post-operative functional status using the Oxford Knee Score (OKS) patient reported outcome tool.

[Response Ends]

**sp.04. Check all the clinical condition/topic areas that apply to your measure, below.**

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Surgery: General*

**[Response Begins]**

Musculoskeletal

Musculoskeletal: Joint Surgery

**[Response Ends]**

**sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.**

**[Response Begins]**

Health and Functional Status: Change

Person-and Family-Centered Care: Person-and Family-Centered Care

**[Response Ends]**

**sp.06. Select one or more target population categories.**

*Select only those target populations which can be stratified in the reporting of the measure's result.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Populations at Risk: Populations at Risk*

**[Response Begins]**

**[Response Ends]**

**sp.07. Select the levels of analysis that apply to your measure.**

*Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- *Clinician: Clinician*
- *Population: Population*

**[Response Begins]**

Clinician: Group/Practice

**[Response Ends]**

**sp.08. Indicate the care settings that apply to your measure.**

*Check ONLY the settings for which the measure is SPECIFIED and TESTED.*

**[Response Begins]**

Outpatient Services

**[Response Ends]**

**sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.**

*Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".*

**[Response Begins]**

<http://mncm.org/wp-content/uploads/2013/04/Total-Knee-Replacement-Data-Collection-Guide-2014-FINAL-2.19.2014.pdf>

**[Response Ends]**

**sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.**

*Attach an excel or csv file; if this poses an issue, [contact staff](#). Provide descriptors for any codes. Use one file with multiple worksheets, if needed.*

**[Response Begins]**

No data dictionary/code table – all information provided in the submission form

**[Response Ends]**

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.12. State the numerator.**

*Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).*

*DO NOT include the rationale for the measure.*

**[Response Begins]**

There is not a traditional numerator for this measure; the measure is calculating the average change in functional status score from pre-operative to post-operative functional status score. The measure is NOT aiming for a numerator target value for a post-operative OKS score.

For example:

The average change in knee function was an increase of 15.9 points one year post-operatively on a 48 point scale.

**[Response Ends]**

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.13. Provide details needed to calculate the numerator.**



*All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.*

*Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

There is not a traditional numerator for this measure; the measure is calculating the average change in functional status score from pre-operative to post-operative functional status score. The measure is NOT aiming for a numerator target value for a post-operative OKS score.

The average change is calculated as follows:

Change is first calculated for each patient and then changed scores are summed and then an average is determined. Measure calculation takes into account those patients that have an improvement and those patients whose function decreases post-operatively. Example below:

Patient Preop OKS Score 1 Year Postop OKS Score Change in OKS Score at 1 Year

Patient A 33 45 12

Patient B 17 39 22

Patient C 16 31 15

Patient D 23 40 17

Patient E 34 42 8

Patient F 10 42 32

Patient G 14 44 30

Patient H 32 44 12

Patient I 19 45 26

Patient J 26 19 -7

Patient K 24 43 19

Patient L 29 34 5

Patient M 23 39 16

Patient N 29 45 16

Patient O 29 45 16

Patient P 34 41 7

Patient Q 11 14 3

Patient R 13 39 26

Patient S 18 45 27

15.9 increase in points on a 48 point scale

**[Response Ends]**

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.14. State the denominator.**

*Brief, narrative description of the target population being measured.*

**[Response Begins]**

Adult patients age and older (no upper age limit) who undergo a primary or revision total knee replacement procedure during a calendar year performance period (e.g. dates of procedure occurring between 1/1/2013 and 12/31/2013) AND have a completed pre-operative and post-operative OKS patient reported outcome assessments.

**[Response Ends]**

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

**sp.15. Provide details needed to calculate the denominator.**

*All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.*

*Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

The initial patient population is adult patients age 18 and older (no upper age limit) who undergo a primary or revision total knee replacement procedure during a calendar year performance period (e.g. dates of procedure occurring between 1/1/2013 and 12/31/2013).

CPT procedure codes: 27445-27447, 27486, 27487

- \* Primary total knee replacement is defined as the first total knee replacement for this particular knee joint.
- \* Revision total knee replacement is defined as the replacement of the previous failed total knee prosthesis with a new prosthesis. Some of the reasons for failure include wear, loosening, infection, fracture, instability, and patient related factors.
- \* Patients with either a primary a revision total knee replacement are included, however the functional status outcome rates will be reported separately (stratified).
- \* Patients with bilateral knee replacements (both knees replaced on the same day, during the same procedure) are included. This would be one procedure based record for submission.
- \* Patients with sequential knee replacements (each knee replaced on a separate day, during a separate procedure) are included. This patient would have two procedure based records, one for each procedure.

Inclusion in the denominator that measures the average change between pre-operative and post-operative functional status requires completion of a patient reported outcome assessment tool (OKS) BOTH pre-operatively (within three months prior to the procedure) AND one year post-operatively (nine to fifteen months after the procedure)

The denominator for calculating the average change in function at a practice level is those patients included in the initial patient population who have both a completed pre-operative and post-operative Oxford Knee Score (OKS) patient reported outcome tool.

**[Response Ends]**

**sp.16. Describe the denominator exclusions.**

*Brief narrative description of exclusions from the target population.*

**[Response Begins]**

There are no denominator exclusions from the initial patient population for this measure.

**[Response Ends]**

**sp.17. Provide details needed to calculate the denominator exclusions.**

*All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.*

**[Response Begins]**

**[Response Ends]**

**sp.18. Provide all information required to stratify the measure results, if necessary.**

*Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.*

**[Response Begins]**

Primary versus revision total knee replacement is the stratification variable for this measure; it is the intent of the measure development group that the outcome rates for this variable are always used and reported separately. As part of the patient level submission of demographic data and PRO tool scores that are submitted to MNMCM's HIPAA secure data portal, a field called Procedure Type is included. Definitions and directions for this field include the following:

Procedure Type:

Enter the type of total knee replacement for this procedure date:

1 = Primary Total Knee Replacement

2 = Revision Total Knee Replacement

This field will be used to stratify results by primary or revision patients.

May use the primary CPT codes to determine the status of primary or revision.

This variable is defined by CPT codes as follows:

Primary Total Knee Replacement Procedures:

CPT Code CPT Procedure Code Description

27445 Arthroplasty, knee hinge prosthesis

27446 Arthroplasty, knee condyle and plateau, medial OR lateral compartment

27447 Arthroplasty, knee condyle and plateau, medial AND lateral compartment with or without patellar resurfacing (total knee arthroplasty)

Revision Total Knee Replacement Procedures:

CPT Code CPT Procedure Code Description

27486 Revision of total knee arthroplasty, with or without allograft, 1 component

27487 Revision of total knee arthroplasty, with or without allograft, femoral and entire tibial component

**[Response Ends]**

**sp.19. Select the risk adjustment type.**

*Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.*

**[Response Begins]**

Statistical risk model

**[Response Ends]**

**sp.20. Select the most relevant type of score.**

*Attachment: If available, please provide a sample report.*

**[Response Begins]**

Continuous variable, e.g. average

**[Response Ends]**

**sp.21. Select the appropriate interpretation of the measure score.**

*Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score*

**[Response Begins]**

Better quality = Score within a defined interval

**[Response Ends]**

**sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.**

*Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.*

**[Response Begins]**

Please also refer to measure flow logic in the data dictionary in S.2.b and flow chart in Appendix A-1

Initial patient population:

Was the patient born on or prior to 01/01/xxxx?

Did the patient undergo a primary or revision total knee replacement procedure between 01/01/2012 to 12/31/2012?

Does the patient have one of the following CPT codes?

27445, 27446, 27447, 27486, 27487

Inclusion in Denominator (has pre-op and post-op OKS)

Valid date in the Preop OKS Date field? No = remove from denominator; Yes continue

Is the Preop OKS Date field within 3 months prior to the procedure? No = remove from denominator; Yes continue

Is there a value in the Preop OKS Score field? Yes = Preop OKS Hold this score for calculation if postop score is present, if No remove from denominator.

Is the 1 Yr Postop OKS Date field within nine to fifteen months after the Date of Procedure? No = remove from denominator; Yes continue.

Is there a value in the 1 Yr Postop OKS Score field? If Yes 1 Yr Post-op OKS Hold this score for calculation, if No remove from denominator. .

For each patient remaining in the denominator calculate the change in function by taking the one year post-op OKS score and subtracting pre-op OKS score. Save this change score.

To calculate the rate of average change in functional status for the practice; average the change in function score.

Example:

Patient Preop OKS Score 1 Year Postop OKS Score Change in OKS Score at 1 Year

Patient A 33 45 12

Patient B 17 39 22  
Patient C 16 31 15  
Patient D 23 40 17  
Patient E 34 42 8  
Patient F 10 42 32  
Patient G 14 44 30  
Patient H 32 44 12  
Patient I 19 45 26  
Patient J 26 19 -7  
Patient K 24 43 19  
Patient L 29 34 5  
Patient M 23 39 16  
Patient N 29 45 16  
Patient O 29 45 16  
Patient P 34 41 7  
Patient Q 11 14 3  
Patient R 13 39 26  
Patient S 18 45 27  
15.9 increase in points on a 48 point scale

**[Response Ends]**

**sp.23. Attach a copy of the instrument (e.g. survey, tool, questionnaire, scale) used as a data source for your measure, if available.**

**[Response Begins]**

**[Response Ends]**

**sp.24. Indicate the responder for your instrument.**

**[Response Begins]**

**[Response Ends]**

**sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.**

**[Response Begins]**

This procedurally based measure is based on the full population of eligible patients; sampling is not used.

The patient reported outcome tool for this measure is the Oxford Knee Score. We have not had issues related to proxy completion for this adult measure. Developer instructions for the tool (Ray Fitzpatrick & Jill Dawson and ISIS Outcomes, ISIS Innovation, Ltd.) indicate that the tool should be completed by the patient.

**[Response Ends]**

**sp.26. Identify whether and how proxy responses are allowed.**

**[Response Begins]**

**[Response Ends]**

**sp.27. Survey/Patient-reported data.**

*Provide instructions for data collection and guidance on minimum response rate. Specify calculation of response rates to be reported with performance measure results.*

**[Response Begins]**

MNCM also calculates rates for tool administration as this measure is dependent on consistent administration of PRO assessment tools to patients. Prior to any use of these outcome measures, we evaluate the rates of:

- \* Rate of pre-operative OKS administered within three months prior to the date of procedure
- \* Rate of post-operative OKS administered within nine to fifteen months post-operatively (one year)
- \* Total and rate of patients who have both pre-operative and post-operative tool administration (denominator)

During the measure development process, the work group anticipated fairly high rates of pre-operative administration and had an expectation that 70% of patients could be captured post-operatively, based on the percentage of patients they were seeing for a one year post-op visit. It is noted that a face-to-face visit is not required for a follow-up OKS assessment; mail or email administration is acceptable.

**[Response Ends]**

**sp.28. Select only the data sources for which the measure is specified.**

**[Response Begins]**

Instrument-Based Data

Paper Medical Records

**[Response Ends]**

**sp.29. Identify the specific data source or data collection instrument.**

*For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.*

**[Response Begins]**

Oxford Knee Score Version 1.0 (User Guide August 2010)

A twelve item patient-reported questionnaire originally developed and validated specifically to assess function and pain in patients undergoing total knee replacement. It is short, reproducible, valid and sensitive to clinically important changes (Dawson et al, 1998).

Items are scored on a 0 to 4 ordered scale with 4 indicating the best outcome (least amount of symptoms). One example of a question: During the past four weeks ... Have you had any trouble getting in and out of a car or using public transportation because of your knee? 4 = no trouble at all, 3 = very little trouble, 2 = moderate trouble, 1 = extreme difficulty and 0 = impossible to do. The tool is scored by simply summing all of the responses to the individual questions. Summary scores range from 0 (worst possible outcome) to 48 (best possible outcome) Time for patient completion is 5 to 10 minutes.

Internal consistency: Cronbach's alpha for the study questionnaire was 0.87 before the operation (n=117) and 0.93 at the six-month follow-up (n=85). All but three items correlated with the total score at r=0.53 (items 6, 8 and 10 r=0.45) at the preoperative assessment (Table 1). After surgery all 12 items correlated with the total score at r=0.51. Cronbach's alpha was not markedly improved by removal of any item from the score. (Dawson et al, 1998).

Reproducibility: In the test-retest sample, the correlation (r=0.92) between the total scores for the questionnaire was high (p<0.0001). No significant change occurred in the distribution of scores between the two assessments for reliability (paired t-test >0.05). (Dawson et al, 1998).

Construct validity: The study questionnaire correlated moderately with both components of the AKS clinical scores before operation (Table 2). There was also significant agreement ( $r>0.5$  to  $r=0.71$ ,  $p<0.01$ ) between the OKS questionnaire and relevant domains of the SF 36 (physical function, role physical, pain and social function), and with both components of the HAQ, (pain VAS and the disability index). (Dawson et al, 1998).

Sensitivity to change: Patients reported a substantial improvement at the six-month follow-up assessment. The effect size (2.19) was larger for the OKS questionnaire than for any of the individual subscales of the SF-36 questionnaire (Table 3), indicating that it could be particularly sensitive to improvements obtained by TKR. The change scores for the TKR questionnaire were significantly greater ( $p<0.0001$ ) for patients who reported the most improvement in their condition. (Dawson et al, 1998).

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]

[Response Ends]

**2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).**

*Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:*

**Current Submission:**

*Updated testing information here.*

**Previous Submission:**

*Testing from the previous submission here.*

[Response Begins]

[Response Ends]

**2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).**

*Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:*

**Current Submission:**

*Updated testing information here.*

**Previous Submission:**

*Testing from the previous submission here.*

[Response Begins]

[Response Ends]

**2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?**

[Response Begins]

[Response Ends]

**2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.**

**Please update the Scientific Acceptability: Validity - Other Threats to Validity section.**

**Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.**

[Response Begins]

[Response Ends]

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the [Submitting Standards webpage](#).
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the [2021 Measure Evaluation Criteria and Guidance](#).

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;



AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

## Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v. \$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Scientific Acceptability sections. For example:

**2021 Submission:**

Updated testing information here.

**2018 Submission:**

Testing from the previous submission here.

**2a.01. Select only the data sources for which the measure is tested.**

[Response Begins]

[Response Ends]

**2a.02. If an existing dataset was used, identify the specific dataset.**

*The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).*

[Response Begins]

[Response Ends]

**2a.03. Provide the dates of the data used in testing.**

*Use the following format: "MM-DD-YYYY - MM-DD-YYYY"*

[Response Begins]

[Response Ends]

**2a.04. Select the levels of analysis for which the measure is tested.**

*Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.*

*Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.*

*Please do not select:*

- Clinician: Clinician
- Population: Population

[Response Begins]

[Response Ends]

**2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).**

*Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.*

[Response Begins]

[Response Ends]

**2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.**

*If there is a minimum case count used for testing, that minimum must be reflected in the specifications.*

[Response Begins]

[Response Ends]

**2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.**

[Response Begins]

[Response Ends]

**2a.08. List the social risk factors that were available and analyzed.**

*For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.*

[Response Begins]

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter “see validity testing section of data elements”; and enter “N/A” for 2a.09 and 2a.10.

**2a.09. Select the level of reliability testing conducted.**

*Choose one or both levels.*

[Response Begins]

[Response Ends]

**2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.**

*Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?**

*For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, [NQF Measure Evaluation Criteria](#)).*

[Response Begins]

[Response Ends]

**2a.12. Interpret the results, in terms of how they demonstrate reliability.**

*(In other words, what do the results mean and what are the norms for the test conducted?)*

[Response Begins]

[Response Ends]

**2b.01. Select the level of validity testing that was conducted.**

[Response Begins]

[Response Ends]

**2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.**

*Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.03. Provide the statistical results from validity testing.**

*Examples may include correlations or t-test results.*

[Response Begins]

[Response Ends]

**2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)**

[Response Begins]

[Response Ends]

**2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.**

*Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.*

[Response Begins]

[Response Ends]

**2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.**

*Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.*

[Response Begins]

[Response Ends]

**2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.**

*In other words, what do the results mean in terms of statistical and meaningful differences?*

[Response Begins]

[Response Ends]

**2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.**

*Describe the steps—do not just name a method; what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.**

*For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).*

[Response Begins]

[Response Ends]

**2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.**

*In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.*

[Response Begins]

[Response Ends]

Note: This item is directed to measures that are risk-adjusted (with or without social risk factors) OR to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eQMs). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

**2b.11. Indicate whether there is more than one set of specifications for this measure.**

[Response Begins]

[Response Ends]

**2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.**

*Describe the steps—do not just name a method. Indicate what statistical analysis was used.*

[Response Begins]

[Response Ends]

**2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.**

*Examples may include correlation, and/or rank order.*

[Response Begins]

[Response Ends]

**2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.**

*In other words, what do the results mean and what are the norms for the test conducted.*

[Response Begins]

[Response Ends]

**2b.15. Indicate whether the measure uses exclusions.**

[Response Begins]

[Response Ends]

**2b.16. Describe the method of testing exclusions and what was tested.**

*Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?*

[Response Begins]

[Response Ends]

**2b.17. Provide the statistical results from testing exclusions.**

*Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.*

[Response Begins]

[Response Ends]

**2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.**

*In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.*

[Response Begins]

[Response Ends]

**2b.19. Check all methods used to address risk factors.**

[Response Begins]

[Response Ends]

**2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.**

[Response Begins]

[Response Ends]

**2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.**

[Response Begins]

[Response Ends]

**2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.**

[Response Begins]

[Response Ends]

**2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.**

*Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of  $p < 0.10$  or other statistical tests; correlation of  $x$  or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).*

[Response Begins]

[Response Ends]

**2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.**

[Response Begins]

[Response Ends]

**2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.**

*Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.*

[Response Begins]

[Response Ends]

**2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.**

*Validation testing should be conducted in a data set that is separate from the one used to develop the model.*

[Response Begins]



[Response Ends]

**2b.27. Provide risk model discrimination statistics.**

*For example, provide c-statistics or R-squared values.*

[Response Begins]

[Response Ends]

**2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).**

[Response Begins]

[Response Ends]

**2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.**

*The preferred file format is .png, but most image formats are acceptable.*

[Response Begins]

[Response Ends]

**2b.30. Provide the results of the risk stratification analysis.**

[Response Begins]

[Response Ends]

**2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).**

*In other words, what do the results mean and what are the norms for the test conducted?*

[Response Begins]

[Response Ends]

**2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.**

*Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.*

[Response Begins]

[Response Ends]

### 3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

---

**3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.**

**[Response Begins]**

Coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)

**[Response Ends]**

**3.02. Detail to what extent the specified data elements are available electronically in defined fields.**

*In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.*

**[Response Begins]**

ALL data elements are in defined fields in electronic health records (EHRs)

**[Response Ends]**

**3.03. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.**

**[Response Begins]**

**[Response Ends]**

**3.04. Describe any efforts to develop an eCQM.**

**[Response Begins]**

**[Response Ends]**

**3.06. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

**[Response Begins]**

**[Response Ends]**

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

**3.07. Detail any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm),**

**Attach the fee schedule here, if applicable.**

**[Response Begins]**

The Oxford Knee Score tool is a proprietary tool; all use requires written permission or license of copyright from Isis Outcomes/ Isis Innovation, Ltd. At [www.isis-innovation.com/outcomes](http://www.isis-innovation.com/outcomes). MNCM obtained permission from Isis Outcomes for the free use of this tool in clinical practice in Minnesota and bordering communities (practices within 50 miles of the state border) for the purpose of capturing patient reported outcomes for this measure. There are no fees associated with submitting patient level data to the MNCM data portal for rate calculation. There are costs to the practices in the collection and extraction of their data for an annual submission of data.

**[Response Ends]**

## 4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

---

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement, in addition to demonstrating performance improvement.

### 4a.01. Check all current uses. For each current use checked, please provide:

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Public Reporting

Regulatory and Accreditation Programs

[Response Ends]

### 4a.02. Check all planned uses.

[Response Begins]

Public reporting

[Response Ends]

### 4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

*For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?*

[Response Begins]

Plans in place for publicly reporting in 2016 on our consumer facing website MN HealthScores at [www.mnhealthscores.org/](http://www.mnhealthscores.org/)

[Response Ends]

### 4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

*A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*

**[Response Begins]**

Planned Use:

\* Included in the MN Department of Health (MDH) Statewide Quality Reporting and Measurement System. Mandatory data collection and reporting under 2008 MN Health Reform Legislation. MNCM was a subcontractor to MDH for measure development exploring the concept total knee replacement.

**[Response Ends]**

**4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.**

*Detail how many and which types of measured entities and/or others were included. If only a sample of measured entities were included, describe the full population and how the sample was selected.*

**[Response Begins]**

**[Response Ends]**

**4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**

**[Response Begins]**

**[Response Ends]**

**4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.**

**[Response Begins]**

**[Response Ends]**

**4a.08. Summarize the feedback obtained from those being measured.**

**[Response Begins]**

**[Response Ends]**

**4a.09. Summarize the feedback obtained from other users.**

**[Response Begins]**

**[Response Ends]**

**4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**

**[Response Begins]**

**[Response Ends]**

**4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**

**[Response Begins]**

**[Response Ends]**

**4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.**

**[Response Begins]**

No negative unintended consequences identified during pilot. As measure moves forward with wider implementation; need to be mindful of potential gaming (cherry-picking); this can be managed through monitoring the pre-operative and post-operative PRO administration rates. One would expect a fairly high pre-operative rate (80's to 90's) that can be incorporated into pre-operative processes and a post-operative capture rate

**[Response Ends]**

**4b.03. Explain any unexpected benefits realized from implementation of this measure.**

**[Response Begins]**

**[Response Ends]**

## 5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

---

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

### 5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

*(Can search and select measures.)*

[Response Begins]

[Response Ends]

### 5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

*(Can search and select measures.)*

[Response Begins]

[Response Ends]

### 5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

[Response Ends]

### 5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]

[Response Ends]

### 5.05. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

[Response Begins]

[Response Ends]

### 5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

*Provide analyses when possible.*

[Response Begins]

[Response Ends]

## Appendix

Supplemental materials may be provided in an appendix.:

## Contact Information

**Measure Steward (Intellectual Property Owner):** MN Community Measurement

**Measure Steward Point of Contact:** Cole, Collette, cole@mncm.org

**Measure Developer if different from Measure Steward:** MN Community Measurement

**Measure Developer Point(s) of Contact:** Cole, Collette, cole@mncm.org



## Additional Information

**1. Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.**

[Response Begins]

[Response Ends]

**2. List the workgroup/panel members' names and organizations.**

*Describe the members' role in measure development.*

[Response Begins]

Marc Swiontkowski, MD Clinical Provider: Surgeon; Chair TRIA Orthopedics  
Andrew Schmidt, MD Clinical Provider: Surgeon HFA & HCMC Clinics  
Scott Anseth, MD Clinical Provider: Surgeon Twin Cities Orthopedics  
Scott Marston, MD Clinical Provider: Surgeon HealthPartners  
Joseph Signorelli, MD Clinical Provider: Surgeon Essentia Health- Orthopedics  
Gregg Strathy, MD Clinical Provider: Surgeon Park Nicollet  
Jeff Temple Clinical Provider: CFNP-OPA Northern Pines Clinic  
Tad Mabry, MD Clinical Provider: Surgeon Mayo Clinic  
Lisa Aker Data Analyst HealthPartners  
Pam York State Agency Minnesota Department of Health  
Sunny Ray QI or Clinic Admin TRIA Orthopedics  
Cara Broich Health Plan Medica  
Collette Pitzen Facilitator/ Measure Dev MNMCM

[Response Ends]

**3. Indicate the year the measure was first released.**

[Response Begins]

[Response Ends]

**4. Indicate the month and year of the most recent revision.**

[Response Begins]

[Response Ends]

**5. Indicate the frequency of review, or an update schedule, for this measure.**

[Response Begins]

Annual review and update

[Response Ends]

**6. Indicate the next scheduled update or review of this measure.**

[Response Begins]

[Response Ends]

**7. Provide a copyright statement, if applicable. Otherwise, indicate "N/A".**

**[Response Begins]**

© MN Community Measurement, 2014. All rights reserved.

Oxford Knee Score (OKS) © Copyright, Isis Innovation Limited 1998.

**[Response Ends]**

**8. State any disclaimers, if applicable. Otherwise, indicate "N/A".**

**[Response Begins]**

**[Response Ends]**

**9. Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".**

**[Response Begins]**

**[Response Ends]**