



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 IM1.1 relates to sub criterion IM1).

Brief Measure Information

NQF #: 3512

De.2. Measure Title: Knee Arthroplasty

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: The Knee Arthroplasty cost measure evaluates clinicians' risk-adjusted cost to Medicare for beneficiaries who receive this procedure. The cost measure score is a clinician's average risk-adjusted cost for the episode group averaged across all episodes attributed to the clinician. This procedural measure includes costs of services that are clinically related to the attributed clinician's role in managing care during the 30 days prior to the clinical event that opens or 'triggers' the episode, through 90 days after the trigger. Beneficiary populations eligible for the Knee Arthroplasty measure include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

IM.1.1. Developer Rationale: An estimated 45 percent of adults in the United States are at risk for developing knee osteoarthritis at some point in their life, and the rate of Medicare beneficiaries undergoing a Knee Arthroplasty to treat it (or other forms of arthritis) has recently increased. From 2000 to 2006, the rate increased by 58 percent from a rate of 55 per 10,000 to a rate of 85 per 10,000 Medicare beneficiaries.[1] Opportunities for improvement include potential for a reduction in readmissions and mitigation of Venous Thromboembolism (VTE), which can occur after a Knee Arthroplasty and result in a significantly more expensive and longer hospital stay. The Knee Arthroplasty episode-based cost measure was recommended for development by an expert clinician committee—the Musculoskeletal Disease Management - Non-Spine Clinical Subcommittee—because of its high impact in terms of patient population and Medicare spending, and the opportunity for incentivizing cost-effective, high-quality clinical care in this area. The Clinical Subcommittee provided extensive, detailed input on this measure.

[1] M. G. Cisternas et al., "Racial Disparities in Total Knee Replacement Among Medicare Enrollees -- United States, 2000-2006. (cover story)," MMWR: Morbidity & Mortality Weekly Report 58, no. 6 (2009).

De.1. Measure Type: Cost/Resource Use

S.5. Data Source: Claims

Enrollment Data

Other

S.3. Level of Analysis: Clinician : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Nov 13, 2019 **Most Recent Endorsement Date:**

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

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

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

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.**IM.1. Opportunity for Improvement****IM.1.1. Briefly explain the rationale for this measure (e.g., the benefits or improvements in performance envisioned by use of this measure)**

An estimated 45 percent of adults in the United States are at risk for developing knee osteoarthritis at some point in their life, and the rate of Medicare beneficiaries undergoing a Knee Arthroplasty to treat it (or other forms of arthritis) has recently increased. From 2000 to 2006, the rate increased by 58 percent from a rate of 55 per 10,000 to a rate of 85 per 10,000 Medicare beneficiaries.[1] Opportunities for improvement include potential for a reduction in readmissions and mitigation of Venous Thromboembolism (VTE), which can occur after a Knee Arthroplasty and result in a significantly more expensive and longer hospital stay. The Knee Arthroplasty episode-based cost measure was recommended for development by an expert clinician committee—the Musculoskeletal Disease Management - Non-Spine Clinical Subcommittee—because of its high impact in terms of patient population and Medicare spending, and the opportunity for incentivizing cost-effective, high-quality clinical care in this area. The Clinical Subcommittee provided extensive, detailed input on this measure.

[1] M. G. Cisternas et al., "Racial Disparities in Total Knee Replacement Among Medicare Enrollees -- United States, 2000-2006. (cover story)," MMWR: Morbidity & Mortality Weekly Report 58, no. 6 (2009).

IM.1.2. Provide performance scores on the measure as specified (current and over time) **at the specified level of analysis.** (This is required for endorsement maintenance. Include mean, stddev, min, max, interquartile range, 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 subcriterion on improvement (U.3.1.) under Usability and Use.

Performance scores are provided for 2,993 clinician group practices (identified by Tax Identification Number [TIN]) and 10,742 practitioners (identified by combination of TIN and National Provider Identifier [NPI]). Clinicians and clinician groups are included if they are attributed 10 or more Knee Arthroplasty episodes, as identified in Medicare Parts A and B claims data, ending from January 1, 2017, to December 31, 2017. Episodes are included from all 50 States and D.C. in the following settings: inpatient (IP) hospitals, hospital outpatient departments (HOPD), ambulatory/office-based care centers, and ambulatory surgical centers (ASC).

TIN Level Scores

- Mean score: \$19,645
- Standard deviation: \$2,181
- Min score: \$12,739
- Max score: \$31,053
- Score IQR: \$2,704
- Score percentiles
 - o 10th: \$17,088
 - o 20th: \$17,825
 - o 30th: \$18,442
 - o 40th: \$18,930
 - o 50th: \$19,490
 - o 60th: \$19,989
 - o 70th: \$20,523
 - o 80th: \$21,233
 - o 90th: \$22,285
- Number of beneficiaries: 237,376

TIN-NPI Level Scores

- Mean score: \$19,052
- Standard deviation: \$2,134
- Min score: \$13,689
- Max score: \$31,138
- Score IQR: \$2,875
- Score percentiles
 - o 10th: \$16,473

- o 20th: \$17,157
- o 30th: \$17,778
- o 40th: \$18,357
- o 50th: \$18,893
- o 60th: \$19,443
- o 70th: \$20,007
- o 80th: \$20,717
- o 90th: \$21,757
- Number of beneficiaries: 227,075

IM.1.3. If no or limited performance data on the measure as specified is reported in IM.1.2., 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.

N/A.

IM.1.4. 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. (This is required for endorsement maintenance. 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 subcriterion on improvement (U.3.1.) under Usability and Use.**

N/A.

IM.1.5. If no or limited data on disparities from the measure as specified is reported in IM.1.4., then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.

N/A.

IM.2. Measure Intent

IM.2.1. Describe intent of the measure and its components/ Rationale (including any citations) for analyzing variation in resource use in this way.

The Knee Arthroplasty measure was developed for use in the Merit-based Incentive Payment System (MIPS) in the Quality Payment Program (QPP), to meet the requirements of MACRA section 101(f). This program is required by law, and aims to achieve high-value care in the Medicare program by measuring clinician performance through four areas: quality, improvement activities, promoting interoperability, and cost. Within the MIPS cost performance category, this measure is intended to provide actionable information to clinicians about their cost performance for this knee replacement procedure, to allow them to make practical changes towards providing high-value, cost effective care.

Rationale for Measuring Cost through Episode-Based Cost Measure vs. All-Cost Measure

The intent of an episode-based cost measure is to capture only clinically related services within the reasonable influence of the attributed clinician, which is a key difference from broad, population-based cost measures such as the MIPS Total Per Capita Cost (TPCC) and Medicare Spending Per Beneficiary (MSPB) measures.

Episode-based cost measures represent the cost to Medicare for the items and services provided to a patient during an episode of care and are meant to inform attributed clinicians about the cost of care within their influence during the episode's timeframe. They represent a clinically cohesive set of medical services rendered to treat a given condition or related to a procedure; services are assigned to an episode only when clinically related to the attributed clinician's role in managing patient care during the episode.

Rationale for Measuring Cost of Knee Arthroplasty

Policymakers contend that an estimated 80 percent of overall health care costs are attributable to decisions made by clinicians.[1] However, these same clinicians are often unaware of how their care decisions influence the overall costs of care. One of the goals for using cost measures is to help inform clinicians on the costs attributable to their decision-making, as well as the total cost of their patient's care. A cost measure offers opportunity for improvement if clinicians can exercise influence on a significant share of costs during the episode, or if lower spending and better care quality can be achieved through changes in clinical practice.

According to the literature and previous feedback received through stakeholder input activities, this measure represents an area where there are significant opportunities for improvement, especially in mitigating costly complications as a result of knee arthroplasty.

Knee arthroplasties account for a sizeable share of Medicare fee-for-service costs among Medicare beneficiaries. Acumen, LLC conducted an empirical analysis of Medicare costs using Medicare administrative claims data and CMS Enrollment Database (EDB) data from August 1, 2015 to July 31, 2016. According to that analysis, more than 241,000 Medicare beneficiaries (0.69 percent of all beneficiaries) underwent a knee arthroplasty, with total annual episode costs accounting for between approximately 0.65 percent to 1.20 percent (\$2.64 to \$4.83 billion) of Medicare Parts A and B fee-for-service spending. A review of the knee arthroplasty literature indicates multiple areas of opportunity for improvement, primary among them reduction of readmissions and mitigation of venous thromboembolism (VTE).

Readmissions for knee arthroplasty are costly and frequent; therefore, lowering readmission rates is a substantial opportunity for potential cost savings in Medicare. According to a 2015 study of 30-day orthopedic readmission rates, approximately 4 to 5 percent of knee arthroplasty procedures for Medicare beneficiaries result in a 30-day readmission.[2] The majority of readmissions following knee arthroplasty are the result of surgical complications and include post-operative surgical site infection, dislocation of a prosthetic joint, periprosthetic fracture, and wound disruption. Readmissions after knee arthroplasty are costly, with 90-day readmissions for surgical complications costing an average of \$28,000 alone, and 90-day readmissions for medical complications costing an average of \$12,000 once outliers were removed.[3]

Reduction of venous thromboembolism (VTE) following knee arthroplasty represents a significant opportunity for reducing cost and improving quality of care. VTE is blood clot in a vein that can have serious consequences such as pulmonary embolism (PE) and requires long-term treatment with blood-thinning medication.[4] Patients with VTE after knee arthroplasties typically have significantly more expensive and longer hospital stays (as well as a greater rate of readmission) relative to patients without VTE, and the condition itself leads to more than 100,000 deaths per year. A 2011 study found that Medicare beneficiaries undergoing knee arthroplasty who subsequently had a VTE had significantly higher rates of mortality, re-hospitalization, and complications, and had risk-adjusted Medicare and total health care costs more than four times greater than patients without VTE.[5]

This measure aims to address these example areas of opportunities for improvement. Since knee arthroplasty is a common procedure in the United States, especially among Medicare beneficiaries, the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs.

[1] Fred, Herbert L. "Cutting the Cost of Health Care: The Physician's Role." Texas Heart Institute Journal, vol. 43, no. 1, 2016, pp. 4 – 6.

[2] Bernatz, James T., Jonathan L. Tueting, and Paul A. Anderson. "Thirty-Day Readmission Rates in Orthopedics: A Systematic Review and Meta-Analysis." [In English]. Plos One 10, no. 4 (Apr 2015).

[3] Clair, Andrew J., Perry J. Evangelista, Claudette M. Lajam, James D. Slover, Joseph A. Bosco, and Richard Iorio. "Cost Analysis of Total Joint Arthroplasty Readmissions in a Bundled Payment Care Improvement Initiative." J Arthroplasty 31, no. 9 (2016): 1862-65.

[4] Baser, Onur, Dylan Supina, Nishan Sengupta, Li Wang, and Louis Kwong. "Impact of postoperative venous thromboembolism on Medicare recipients undergoing total hip replacement or total knee replacement surgery." American Journal of Health-System Pharmacy 67, no. 17 (2010): 1438-1445.

[5] Baser, Onur, Dylan Supina, Nishan Sengupta, Li Wang, and Louis Kwong. "Clinical and cost outcomes of venous thromboembolism in Medicare patients undergoing total hip replacement or total knee replacement surgery." Current medical research and opinion 27, no. 2 (2011): 423-429.

Rationale for Use of Claims Data to Measure Cost

- The use of claims data for episode-based cost measures for MIPS is required by MACRA section 101(f).
- There is no additional submission burden, as clinicians must already submit claims for reimbursement.
- Using Medicare Parts A and B claims data allows CMS to evaluate TIN and TIN-NPI cost across all conditions and procedures, resulting in a comprehensive set of data on knee arthroplasty cost performance.
- Additionally, the wide reach of Medicare claims data maximizes the impact of the measure, ensuring that the most TINs and TIN-NPIs benefit from the information provided on knee arthroplasties.

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.

Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area (check all the areas that apply):

De.6. Non-Condition Specific (check all the areas that apply):

De.7. Care Setting (Select all the settings for which the measure is specified and tested):

S.1. Measure-specific Web Page (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.)

<https://qpp.cms.gov/about/resource-library>. Scroll to “Full Resource Library” to download the Cost Measure Information Form and Measure Code List, or see S.7.2a Construction Logic Attachment for additional details and specific links.

S.2. Type of resource use measure (Select the most relevant)

[Per episode](#)

S.3. Level of Analysis (Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED):

[Clinician : Group/Practice](#), [Clinician : Individual](#)

S.4. Target Population Category (Check all the populations for which the measure is specified and tested if any):

S.5. Data Source (Check ONLY the sources for which the measure is SPECIFIED AND TESTED).

If other, please describe in S.5.1.

[Claims](#)

[Enrollment Data](#)

[Other](#)

S.5.1. Data Source or Collection Instrument (Identify the specific data source or data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.)

The Knee Arthroplasty measure uses Medicare Part A and Part B claims data, which is maintained by CMS. Part A and B claims data are used to build episodes of care, calculate episode costs, and construct risk adjusters. Data from the Medicare Enrollment Database (EDB) are used to determine beneficiary-level exclusions and supplemental risk adjusters, specifically Medicare Parts A, B, and C enrollment; primary payer; disability status; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates. The risk adjustment model also accounts for expected differences in payment for services provided to beneficiaries in long-term care, and that information comes from the Minimum Data Set (MDS). The MDS is used to create the Long Term Care Indicator variable in risk adjustment.

For measure testing, data from the American Census, American Community Survey (ACS), and Common Medicare Enrollment (CME) are used in the analyses evaluating social risk factors in risk adjustment.

S.5.2. Data Source or Collection Instrument Reference (available at measure-specific Web page URL identified in S.1 OR in the file attached here) (Save file as: S_5_2_DataSourceReference)

<SamplingMethodologySpecificDataSourceAttachment nodeType="0">S_5_2_DataSourceReference-636824808585196896.docx

S.6. Data Dictionary or Code Table *(Please provide a web page URL or attachment if exceeds 2 pages. NQF strongly prefers URLs. Attach documents only if they are not available on a web page.)*

Data Dictionary:

URL: The Research Data Assistance Center (ResDAC) maintains the Medicare claims data dictionary available here: <http://www.resdac.org/cms-data/filefamily/Medicare-Claims> CMS maintains the Medicare Enrollment Database and data dictionary: edbonline@cms.hhs.gov

Please supply the username and password:

Attachment:

Code Table:

URL:

Please supply the username and password:

Attachment: 2019_01_07_testing_form_appendix_knee_arthro.xlsx

Construction Logic

S.7.1. Brief Description of Construction Logic

If applicable, summarize the general approach or methodology to the measure construction. This is most relevant to measures that are part of or rely on the execution of a measure system or applies to multiple measures.

The Knee Arthroplasty measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare averaged across the episodes attributed to a clinician or clinician group. This is then multiplied by the national average observed episode cost to generate a dollar figure. The measure can be calculated for an individual TIN-NPI (clinician) or a TIN (clinician group practice).

A Knee Arthroplasty episode is a unit or specific instance of the measure for a given clinician or clinician group and beneficiary that can then be aggregated to assess a clinician's performance across all their episodes. The episode is triggered or opened by Current Procedural Terminology / Healthcare Common Procedure Coding System (CPT/HCPCS) codes, and includes certain services in Medicare Parts A and B claims related to the procedure in the period.

The cost measure numerator is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all Knee Arthroplasty episodes attributed to a clinician or clinician group. This sum is then multiplied by the national average observed episode cost to generate a dollar figure.

The cost measure denominator is the total number of episodes from Knee Arthroplasty episode group attributed to a clinician or clinician group within a performance period (i.e., MIPS performance year).

Cost figures are standardized to remove the effect of differences in Medicare payment among health care providers that are the result of differences in regional health care provider expenses measured by hospital wage indexes and geographic price cost indexes (GPCIs) or other payment adjustments such as those for teaching hospitals. This standardization is intended to isolate cost differences that result from healthcare delivery choices, allowing for more accurate resource use comparisons between health care providers.

S.7.2. Construction Logic *(Detail logic steps used to cluster, group or assign claims beyond those associated with the measure's clinical logic.)*

Step 1. Trigger and Define an Episode

Knee Arthroplasty episodes are defined by Current Procedural Terminology / Healthcare Common Procedure Coding System (CPT/HCPCS) codes on Part B Physician/Supplier (Carrier) claims that open, or trigger, an episode.

The steps for defining an episode for the Knee Arthroplasty episode group are as follows:

- Identify Part B Physician/Supplier claim lines with positive standardized payment that have a trigger code.
- Trigger an episode if all the following conditions are met for an identified Part B Physician/Supplier claim line:
 - o It was billed by a clinician of a specialty that is eligible for MIPS.
 - o It is the highest cost claim line across any Knee Arthroplasty trigger code billed for the beneficiary on that day. If multiple trigger Part B Physician/Supplier claim lines occur on different days within a concurrent IP stay, an episode will be triggered by the trigger Part B Physician/Supplier claim line with the earliest expense date during the IP stay.

- o It does not have a post-operative modifier code.[1]
 - Establish the episode window as follows:
 - o Establish the episode trigger date as the date of admission if an IP stay with a relevant DRG concurrent with the trigger is found, otherwise the expense date of the trigger code.
 - o Establish the episode start date as 30 days prior to the episode trigger date.
 - o Establish the episode end date as 90 days after the episode trigger date.
 - Define trigger exclusions based on information available at the time of the trigger, if applicable.
- Once a Knee Arthroplasty episode is triggered, the episode is placed into one of the episode sub-groups to enable meaningful clinical comparisons. This cost measure has three sub-groups:
- o Partial Knee / Unilateral
 - o Total Knee / Bilateral
 - o Total Knee / Unilateral

Step 2. Attribute Episodes to a Clinician

Once an episode has been triggered and defined, it is attributed to one or more clinicians of a specialty that is eligible for MIPS. Clinicians are identified by Taxpayer Identification Number (TIN) and National Provider Identifier (NPI) pairs (TIN-NPI), and clinician groups are identified by TIN. Only clinicians of a specialty that is eligible for MIPS or clinician groups where the triggering clinician is of a specialty that is eligible for MIPS are attributed episodes.

The steps for attributing a Knee Arthroplasty episode are as follows:

- Identify claim lines with positive standardized payment for any trigger codes that occur on the episode trigger day.
- Designate a TIN-NPI as a main clinician if the following conditions are met:
 - o No assistant modifier code is found on one or more claim lines billed by the clinician.
 - o No exclusion modifier code is found on the same claim line.
- Designate a TIN-NPI as an assistant clinician if the following conditions are met:
 - o The TIN-NPI was not designated as a main clinician.
 - o An assistant modifier code is found.
 - o No exclusion modifier code is found.
- Attribute an episode to any TIN-NPI designated as a main or assistant clinician.
- Attribute episodes to the TIN by aggregating all episodes attributed to NPIs that bill to that TIN. If the same episode is attributed to more than one NPI within a TIN, the episode is attributed only once to that TIN.

Step 3. Assign Costs of Services to an Episode and Calculate Total Observed Episode Cost

For the Knee Arthroplasty episode group, only services performed in the following service categories are considered for assignment to the episode costs:

- o Emergency Department (ED)
- o Outpatient (OP) Facility and Clinician Services
- o Long Term Care Hospital (LTCH) - Medical
- o LTCH - Surgical
- o IP - Medical
- o IP - Surgical
- o Inpatient Rehabilitation Facility (IRF) - Medical
- o Durable Medical Equipment, Prosthetics, Orthotics, and Supplies (DME)
- o Home Health (HH)

Service assignment rules may be modified based on the service category in which the service is performed, as listed above. Service assignment rules may also vary based on (i) additional criteria determined by other diagnosis, procedure, or billing codes appearing alongside the service code, or (ii) the specific timing of the service. Services may be assigned to the episode based on the following additional criteria:

Services may be assigned to the episode based on the following additional criteria:

- o Service code alone
- o Service code in combination with other diagnosis, procedure, or billing codes such as:
 - ? The first three digits of the International Classification of Diseases – Tenth Revision diagnosis code (3-digit ICD-10 DGN)
 - ? The full ICD-10 DGN
 - ? Additional service information

Services may be assigned only with specific timing:

- o Services may be assigned based on whether or not the service and/or diagnosis is newly occurring.
- o Services may be assigned only if they occur within a particular number of days from the trigger within the episode window, and services may be assigned for a period shorter than the full duration of the episode window.

The steps for assigning costs are as follows:

- Identify all services on claims with positive standardized payment that occur within the episode window.
- Assign identified services to the episode based on the types of service assignment rules described above.
- Assign skilled nursing facility (SNF) claims based on the following:
 - o Identify SNF claims for which both (i) the SNF claim's qualifying IP stay is the IP stay during which the trigger occurs, if an IP stay is found, and (ii) the SNF claim occurs during the episode window.
 - o For those identified SNF claims, assign the percentage of the claim amount proportional to the portion of the SNF claim that overlaps with the episode window.
- Sum standardized Medicare allowed amounts for all claims assigned to each episode to obtain the standardized total observed episode cost.

Step 4. Exclude Episodes

The steps for episode exclusion are as follows:

- Exclude episodes from measure calculation if:
 - o The beneficiary has a primary payer other than Medicare for any time overlapping the episode window or 120-day lookback period prior to the trigger day.
 - o The beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or was enrolled in Part C for any part of the lookback plus episode window.
 - o No main clinician is attributed the episode.
 - o The beneficiary's date of birth is missing.
 - o The beneficiary's death date occurred before the episode ended.
 - o The episode trigger claim was not performed in an ambulatory/office-based care, IP hospital, OP hospital, or ASC setting based on its place of service.
 - o The IP facility is not a short-term stay acute hospital as defined by subsection (d) when an IP stay concurrent with the trigger is found.[2]
- Apply measure-specific exclusions, which check the beneficiary's Medicare claims history for certain billing codes (as specified in the Measure Codes List file) that indicate the presence of a particular procedure, condition, or characteristic.

Step 5. Estimate Expected Costs through Risk Adjustment

Steps for defining risk adjustment variables and estimating the risk adjustment model are as follows:

- Define HCC and episode group-specific risk adjustors using service and diagnosis information found on the beneficiary's Medicare claims history in the 120-day period prior to the episode trigger day for certain billing codes that indicate the presence of a procedure, condition, or characteristic.
- Define other risk adjustors that rely upon Medicare beneficiary enrollment and assessment data as follows:
 - o Identify beneficiaries who are originally "Disabled without end-stage renal disease (ESRD)" or "Disabled with ESRD" using the original reason for joining Medicare field in the Medicare beneficiary enrollment database.
 - o Identify beneficiaries with ESRD if their enrollment indicates ESRD coverage, ESRD dialysis, or kidney transplant in the Medicare beneficiary enrollment database in the lookback period.
 - o Identify beneficiaries who have spent at least 90 days in a long-term care institution without having been discharged to the community for 14 days, based on MDS assessment data.
- Drop risk adjustors that are defined for less than 15 episodes nationally for each sub-group to avoid using very small samples.
- Categorize beneficiaries into age ranges using their date of birth information in the Medicare beneficiary enrollment database. If an age range has a cell count less than 15, collapse this with the next adjacent higher age range category.
- Include the MS-DRG of the episode's trigger IP stay, if an IP stay is found, as a categorical risk adjustor.
- Run an ordinary least squares (OLS) regression model to estimate the relationship between all the risk adjustment variables and the dependent variable, the standardized observed episode cost, to obtain the risk-adjusted expected episode cost. A separate OLS regression is run for each episode sub-group nationally.
- Winsorize [3] expected costs as follows.
 - o Assign the value of the 0.5th percentile to all expected episode costs below the 0.5th percentile.
 - o Renormalize[4] values by multiplying each episode's winsorized expected cost by the sub-group's average expected cost, and dividing the resultant value by the sub-group's average winsorized expected cost.

- Exclude [5] episodes with outliers as follows. This step is performed separately for each sub-group.
 - o Calculate each episode's residual as the difference between the re-normalized, winsorized expected cost computed above and the observed cost.
 - o Exclude episodes with residuals below the 1st percentile or above the 99th percentile of the residual distribution.
 - o Renormalize the resultant expected cost values by multiplying each episode's winsorized expected costs after excluding outliers by the sub-group's average standardized observed cost across all episodes originally in the risk adjustment model, and dividing by the sub-group's average winsorized expected cost after excluding outliers.

Step 6. Calculate Measure Scores

- Measure scores are calculated for a TIN or TIN-NPI as follows:
- Calculate the ratio of observed to expected episode cost for each episode attributed to the clinician/clinician group.
- Calculate the average ratio of observed to expected episode cost across the total number of episodes attributed to the clinician/clinician group.
- Multiply the average ratio of observed to expected episode cost by the national average observed episode cost to generate a dollar figure representing risk-adjusted average episode cost.

[1] Post-operative modifier codes indicate that a clinician billing the service was not involved in the main procedure but was involved in the post-operative care for that procedure, and as such the post-operative clinician would not be responsible for the trigger.

[2] Only stays at IP facilities that are a short-term stay acute hospital as defined by subsection (d) will be included. Subsection (d) hospitals are hospitals in the 50 states and D.C. other than: psychiatric hospitals, rehabilitation hospitals, hospitals whose inpatients are predominantly under 18 years old, hospitals whose average inpatient length of stay exceeds 25 days, and hospitals involved extensively in treatment for or research on cancer.

[3] Winsorization aims to limit the effects of extreme values on expected costs. Winsorization is a statistical transformation that limits extreme values in data to reduce the effect of possible outliers. Winsorization of the lower end of the distribution (i.e., bottom coding) involves setting extremely low predicted values below a predetermined limit to be equal to that predetermined limit.

[4] Renormalization is performed after adjustments are made to the episode's expected cost, such as bottom-coding or residual outlier exclusion. This process multiplies the adjusted values by a scalar ratio to ensure that the resulting average is equal to the average of the original value.

[5] This step excludes episodes based on outlier residual values from the calculation and renormalizes the resultant values to maintain a consistent average episode cost level.

S.7.2a. CONSTRUCTION LOGIC ATTACHMENT or URL: If needed, attach supplemental documentation (Save file as:

S_7_2_Construction_Logic). All fields of the submission form that are supplemented within the attachment must include a summary of important information included in the attachment and its intended purpose, including any references to page numbers, tables, text, etc.

URL:

Please supply the username and password:

Attachment: S_7_2_Construction_Logic-636927599486395042.docx

S.7.3. Concurrency of clinical events, measure redundancy or overlap, disease interactions (Detail the method used for identifying concurrent clinical events, how to manage them, and provide the rationale for this methodology.)

To identify and manage concurrent clinical events, the measure takes into consideration the occurrence of the same procedure on the other knee when it occurs within certain periods within the episode window. The measure provides specifications to account for a procedure that is performed on one side only (unilateral), both sides on the same day (same day bilateral), or both sides on different days within short succession (staged bilateral).

The risk adjustment methodology accounts for laterality in the following way:

- o Establish a unilateral sub-group if the procedure is performed on only one side only during the trigger event.

- o Establish a bilateral sub-group if one of the following conditions is true:
- o The procedure is performed on both sides during the trigger event.
- o A procedure on one side (identified by the sole presence of right [or left] modifier code) is followed by the opposite side procedure within 90 days. In these staged laterality events, the two episodes will be combined to one.

Using this methodology to establish sub-groups, the measure accounts for differences in episode cost for unilateral and bilateral procedures. The rationale for having sub-groups based on unilateral and bilateral surgery is to account for some services (e.g., preoperative exams, testing, and rehabilitation) that may be applied to a second surgery performed in close succession. The care pathways of bilateral procedures are also different from those for a unilateral procedure, such as the need for more intensive rehabilitation.

This measure is designed to allow episodes to overlap with other episodes: overlapping episodes are different episodes that are triggered for the same patient with overlapping episode windows. The advantage of this is that each episode can reflect attributed clinicians' different roles in providing care services throughout a patient's care trajectory. For example, a patient could have a Knee Arthroplasty episode triggered when the attributed clinician performs the procedure, and 80 days later be admitted to hospital for pneumonia unrelated to the knee surgery, triggering an episode for a different cost measure that is attributed to the hospitalist providing care for pneumonia. Each episode includes only the cost of assigned services (i.e., those that are within the reasonable influence of the attributed clinician) to reflect each attributed clinician's role. In addition, costs are not double counted as the measure calculation is based on the ratio of observed over expected spending for each episode, then averaged across all of an attributed clinician's episodes.

The measure accounts for disease interactions through its risk adjustment model based on the CMS Hierarchical Condition Category Version 22 (CMS-HCC V22) 2016 model. In addition to the HCCs, the model includes disease interactions (e.g., Cancer * Immune Disorders). Further details about the risk adjustment model and disease interaction terms are included in Section S.8.6.

S.7.4. Complementary services *(Detail how complementary services have been linked to the measure and provide rationale for this methodology.)*

This measure includes the cost of services that are clinically related to the procedure for knee arthroplasty. The rationale for only including specific costs is to ensure that the attributed clinician is evaluated only on his or her performance on services over which they have reasonable influence. For instance, the cost of pre-operative x-ray imaging for knee surgery is included in a clinician's episode cost if it occurs any time during the pre-trigger period.

These services that are assigned to the measure have been identified as being related to the procedure and within the influence of the attributed clinician through consideration of detailed input from clinician experts and broader feedback from stakeholders from the clinician community. Specifically, a Musculoskeletal Disease Management – Non-Spine Clinical Subcommittee was convened from May 2017 to January 2018 to discuss and provide detailed recommendations on aspects of measure construction, including the services to be included in this measure. This Subcommittee was composed of 28 clinician experts affiliated with 27 specialty societies.

The members reviewed analyses of the utilization and timing of all Medicare Parts A and B services in broad timeframes extending before and after the episode trigger to provide recommendations on the services and associated conditions for including these as part of the episode costs. Conditions could include requiring additional codes to be present on services to ensure clinical relevance, assigning for a shorter timeframe within the overall episode window, or assigning only where a diagnosis that is part of the trigger logic is newly occurring. The draft measure was field tested from October to November 2017: during this time, stakeholders reviewed the measure specifications, including a list of assigned services and associated logic conditions, field test reports containing details of attributed clinician performance, and supplemental documentation. Over 65,000 TIN and TIN-NPI field test reports were available during this time for review and feedback.

During field testing, a National Summary Data Report, later updated to include reliability analyses, was posted along with the measure specifications:

- National Summary Data Report (July 2018) – this document contains summary data about the Knee Arthroplasty cost measure, along with other episode-based cost measures. These summary statistics supplement the testing analyses contained in this submission: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2017-field-test-materials.zip>, filename: 2018-07-12-national-summary-data-report.pdf

Stakeholder feedback gathered during field testing was summarized into the Field Testing Feedback Summary Report:

- Field Testing Feedback Summary Report (June 2018) – this document summarizes the feedback received during a stakeholder feedback period during measure development. The Knee Arthroplasty cost measure has been developed with extensive

input from the clinician community: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2018-field-testing-feedback-summary-report.pdf>

S.7.5. Clinical hierarchies *(Detail the hierarchy of codes or condition groups used and provide rationale for this methodology.)*

Clinical hierarchies are embedded in the risk adjustment model. The risk adjustment model includes variables from the CMS-HCC V22 2016 Risk Adjustment Model, as well as other standard risk adjustors (e.g., beneficiary age brackets using information in the Medicare beneficiary enrollment database) and disease interaction terms. The model also includes variables specific to this cost measure, identified through the incorporation of detailed clinical input. These variables include conditions which may influence the episode cost and risk of complication, for example; osteoporosis, post-infectious osteoarthritis, or psoriatic arthritis, amongst others.

The CMS-HCC V22 model uses 79 Hierarchical Condition Category (HCC) indicators derived from the beneficiary's claims in the period 120 days prior to the episode trigger day. Other risk adjustors are originally "Disabled without end-stage renal disease (ESRD)" or "Disabled with ESRD" using the original reason for joining Medicare in the Medicare beneficiary enrollment database. The risk adjustment model also uses an indicator for beneficiaries identified as having had recent need of long-term care (90 days in a long-term care institution without having been discharged to community for 14 days) using MDS assessment data. Additional information about the risk adjustment model is included in Section S.8.6.

The Knee Arthroplasty episode group includes all services identified as being clinically relevant to this procedure. There are logic rules to determine when and what conditions each particular service will be assigned, as detailed in the Measure Codes List file (see Section S.1 for URL).

S.7.6. Missing Data *(Detail steps associated with missing data and provide rationale for this methodology (e.g., any statistical techniques to impute missing data))*

:

All the data used to calculate the Knee Arthroplasty cost measure are included on Medicare claims data. The data fields used to calculate measure (e.g., payment amounts, diagnosis and procedure codes, etc.) are included in all Medicare claims because clinicians only receive payments for complete claims. Additional information regarding the reliability of diagnostic information on claims is available on the Testing Form in Section 2a2.2.

We have complete data for each beneficiary who opens an episode by receiving a triggering service, since beneficiaries are excluded if they are not continuously enrolled in only Medicare Parts A and B or if Medicare is not the primary payer during an episode. This ensures that we have all claims data for beneficiaries included in the Knee Arthroplasty cost measure.

S.7.7. Resource Use Service Categories (Units) *(Select all categories that apply)*

Inpatient services: Inpatient facility services

Inpatient services: Evaluation and management

Inpatient services: Procedures and surgeries

Inpatient services: Imaging and diagnostic

Inpatient services: Lab services

Inpatient services: Admissions/discharges

Other inpatient services

Ambulatory services: Outpatient facility services

Ambulatory services: Emergency Department

Ambulatory services: Evaluation and management

Ambulatory services: Procedures and surgeries

Ambulatory services: Imaging and diagnostic

Ambulatory services: Lab services

Other ambulatory services

Durable Medical Equipment (DME)

Other services not listed

See Measure Codes List

See Measure Codes List

See Measure Codes List

S.7.8. Identification of Resource Use Service Categories (Units)

(For each of the resource use service categories selected above, provide the rationale for their selection and detail the method or algorithms to identify resource units, including codes, logic and definitions.)

The Knee Arthroplasty measure assesses the standardized allowed amounts of services performed by clinicians and other healthcare providers during an episode, which includes all assigned services from Part A and Part B Medicare claims that occur within the time period 30 days prior to the episode trigger through 90 days after the trigger.

The assigned services for this measure are within the following service categories: Outpatient Facilities and Clinician Services, Inpatient Facilities, Emergency Room, Long Term Care Hospital, Inpatient Rehabilitation Facility, Durable Medical Equipment, Prosthetics, Orthotics, and Supplies (DMEPOS), and Home Health. The codes to identify these services (e.g., CPT/HCPCS, Rehabilitation Impairment Category (RIC), DRGs, E&Ms, and Revenue Center) are contained in the Measure Codes List file (see Section S.1), along with the logic conditions for assigning these services.

S.7.8a. If needed, provide supplemental resource use service category specifications in either URL (preferred) or as an attachment (Save file as S.7.8a_RU_Service_Categories):

URL:

Please supply the username and password:

Attachment:

Clinical Logic

S.8.1. Brief Description of Clinical Logic (Briefly describe your clinical logic approach including clinical topic area, whether or not your account for comorbid and interactions, clinical hierarchies, clinical severity levels and concurrency of clinical events.)

This measure aims to provide actionable information to clinicians performing a knee arthroplasty about their resource use within the overall goal of enabling clinicians to provide cost-effective and high-quality care. The clinical logic is constructed to achieve this objective.

Clinical Topic Area: Knee Replacement

Comorbidity and Interactions: The risk adjustment model includes a series of interaction terms between comorbidities and applies a variant of the CMS-HCC risk adjustment model with additional risk adjustors specific to this procedure to capture patient comorbidities.

Clinical Hierarchies: Clinical hierarchies are embedded in the risk adjustment model. See section S.7.5. for further details.

Clinical Severity Levels: The measure has sub-groups to account for the different level of severity for bilateral and unilateral procedures. It also risk adjusts for the MS-DRG where the procedure occurs in an inpatient setting, accounting for medical severity levels.

Concurrency of Clinical Events: The measure spans the period from 30 days prior to episode trigger to 90 days after the episode trigger. Services that are clinically related to the procedure and within the reasonable influence of the attributed clinician within this period of time are included in the episode. The measure accounts for unilateral and bilateral procedures. See Section S.7.3. and

S.7.4. for further details.

S.8.2. Clinical Logic *(Detail any clustering and the assignment of codes, including the grouping methodology, the assignment algorithm, and relevant codes for these methodologies.)*

The Knee Arthroplasty measure includes the cost of follow-up services and those that result as a consequence of care, such as preventable complications, using a service assignment algorithm.

Grouping methodology and assignment algorithm: The Knee Arthroplasty cost measure evaluates resource use through the unit of episodes of care. The cost measure episodes are constructed by including select Medicare Part A and Part B claims (assigned services) which occur during the episode window, defined as 30 days prior to the episode trigger and 90 days after the trigger. The episode triggers and assigned services are contained in the Measure Codes List file (see Section S.1. for details), along with risk adjusters, sub-groups, and exclusions.

Details about the measure exclusions are in Section S.9.1.

Cost Calculation: The cost measure amount includes the cost of assigned services performed by clinicians and other providers during the episode window. The cost measure is calculated as the sum of the ratios of observed to expected costs, multiplied by the national average observed episode cost to generate a dollar figure, and then divided by total number of episodes from the episode group attributed to a clinician. All costs are payment standardized to control for geographic variation in Medicare reimbursement rates. The measure is risk adjusted to account for age and severity of illness. Expected costs are estimated through risk adjustment by using an ordinary least squares regression model. More details about the risk adjustment model are described in Section S.7.5.

S.8.3. Evidence to Support Clinical Logic Described in S.8.2 *Describe the rationale, citing evidence to support the grouping of clinical conditions in the measurement population(s) and the intent of the measure (as described in IM3)*

The clinical logic used in the Knee Arthroplasty measure is informed by literature and stakeholder feedback.

A study notes that policymakers contend that an estimated 80 percent of overall health care costs are attributable to decisions made by clinicians (Fred, 2016). However, these same clinicians are often unaware of how their care decisions influence the overall costs of care. One of the goals of the use of cost measures in general is to help inform clinicians on the costs for which they are directly responsible, as well as the total cost of their patient's care. A cost measure exhibits the opportunity for improvement if clinicians can exercise influence on a significant share of costs during the episode, or if lower spending and better care quality can be made through changes in clinical practice.

Knee arthroplasties account for a sizeable share of Medicare fee-for-service costs among Medicare beneficiaries. According to an empirical analysis of Medicare claims data conducted by Acumen, LLC using Medicare administrative claims data and CMS Enrollment Database (EDB) data spanning from August 1, 2015 to July 31, 2016, more than 241,000 Medicare beneficiaries (0.69 percent of all beneficiaries) underwent a knee arthroplasty, with total annual episode costs accounting for between approximately 0.65 percent to 1.20 percent (\$2.64 to \$4.83 billion) of Medicare Parts A and B fee-for-service spending. A review of the knee arthroplasty literature indicates multiple areas of opportunity for improvement, primary among them reduction of readmissions and mitigation of venous thromboembolism (VTE).

Readmissions for knee arthroplasty are costly and frequent; therefore, lowering readmission rates is a substantial opportunity for potential cost savings in Medicare. According to a 2015 study of 30-day orthopedic readmission rates, approximately 4 to 5 percent of knee arthroplasty procedures for Medicare beneficiaries result in a 30-day readmission (Bernatz et al., 2015). The majority of readmissions following knee arthroplasty are the result of surgical complications and include post-operative surgical site infection, dislocation of a prosthetic joint, periprosthetic fracture, and wound disruption. Readmissions after knee arthroplasty are costly, with 90-day readmissions for surgical complications identified costing an average of \$28,000 alone, and 90-day readmissions for medical complications costing an average of \$12,000 once outliers were removed (Clair et al., 2016).

The measure was designed to incorporate extensive expert clinician input into each component of the measure to ensure that it achieves the goal of providing actionable information to clinicians for their performance of a procedure on a cohesive patient cohort. The measure was developed to meet the requirements of MACRA section 101(f) to create episode-based cost measures. It aligns with CMS meaningful measure area of 'patient-focused episode of care' within the overall quality priority of 'Make Care Affordable'. The measure includes services that are clinically related to the procedure and within the reasonable influence of the attributed clinician. By including services after the procedure, it aims to improve care coordination throughout a patient's care

trajectory.

Bernatz, James T., Jonathan L. Tueting, and Paul A. Anderson. "Thirty-Day Readmission Rates in Orthopedics: A Systematic Review and Meta-Analysis." *Plos One* 10, no. 4 (2015). doi:10.1371/journal.pone.0123593.

Clair, Andrew J., Perry J. Evangelista, Claudette M. Lajam, James D. Slover, Joseph A. Bosco, and Richard Iorio. "Cost Analysis of Total Joint Arthroplasty Readmissions in a Bundled Payment Care Improvement Initiative." *J Arthroplasty* 31, no. 9 (2016): 1862-65.

S.8.3a. CLINICAL LOGIC ATTACHMENT or URL: If needed, attach supplemental documentation (Save file as: S_8_3a_Clinical_Logic). All fields of the submission form that are supplemented within the attachment must include a summary of important information included in the attachment and its intended purpose, including any references to page numbers, tables, text, etc.

URL:

Please supply the username and password:

Attachment: 2018-12-21-codes-list-knee-arthro.xlsx

S.8.4. Measure Trigger and End mechanisms (Detail the measure's trigger and end mechanisms and provide rationale for this methodology)

The detailed steps for triggering Knee Arthroplasty episodes are in Section S.7.2. The advantage of the simplicity in opening episodes this way is to ensure that clinicians know at the time of providing the service that an episode has been triggered. This helps meet the goal of the measure to provide actionable information to clinicians.

Additional conditions must be met to trigger an episode. The Knee Arthroplasty cost measure can be triggered in the following settings: inpatient (IP) hospitals, hospital outpatient departments (HOPD), ambulatory/office-based care centers, and ambulatory surgical centers (ASC).

The Knee Arthroplasty episode window is defined as follows:

- Episode trigger date: expense date of trigger code
- Episode start date: 30 days prior to episode trigger date
- Episode end date: 90 days after episode trigger date

As discussed in Section S.7.3., staged bilateral episodes occurring within 90 days of the initial procedure are collapsed into one episode. The episode's trigger claim has a right (or left) modifier code followed by another Knee Arthroplasty episode with an opposite left (or right) modifier code on a subsequent trigger claim within 90 days. This ensures that assigned services for both procedures are captured.

The conditions to trigger episodes and the duration of the episode window were established with input from clinician experts in consideration of the goals of the measure to provide actionable information to clinicians about their resource use for a comparable patient cohort. An initial Draft List of Episode Groups and Trigger Codes was posted in December 2016 incorporating input from a Clinical Committee of more than 70 clinicians from over 50 professional societies. Feedback from a four-month public comment period on that posting was summarized and shared with the Musculoskeletal Disease Management – Non-Spine Clinical Subcommittee who used the information from the draft list as a starting point and took feedback into consideration along with analyses to help inform discussions, such as the frequency of services over a period of time extending from the trigger date. This measure was field tested in 2017, as discussed further in Section S.7.4. The Clinical Subcommittee took field testing feedback into consideration in making refinements to the measure, including feedback on episode triggers and episode window length.

S.8.5. Clinical severity levels (Detail the method used for assigning severity level and provide rationale for this methodology)

Clinical severity levels are embedded in the risk adjustment model, as described in Section S.7.5.

S.8.6. Comorbid and interactions (Detail the treatment of co-morbidities and disease interactions and provide rationale for this

methodology.)

The Knee Arthroplasty cost measure accounts for comorbid conditions and interactions by broadly following the CMS- Hierarchical Condition Categories (HCCs) risk-adjustment methodology, which is derived from Medicare Part A and B claims and is used in the Medicare Advantage (MA) program. Diagnosis codes on claims that occur during the 120-day period prior to the episode trigger date are used to create HCC indicators. Episodes where the beneficiary is not enrolled in both Medicare Part A and Medicare Part B for the 120 days prior to the episode are excluded because information on comorbidities for these beneficiaries will be incomplete. When applying the CMS-HCC framework to the measure, expected costs are determined by the risk adjustment model separately for each sub-group, which allows the effect of beneficiary health status and demographics on episode spending levels to vary by the sub-groups which discern between total or partial repair and laterality. This cost measure accounts for comorbid interactions by incorporating a number of health status interactions as currently used within the CMS-HCC model. The model includes paired-condition interactions (e.g., chronic obstructive pulmonary disease (COPD) and congestive heart failure (CHF)) and interactions between conditions and disability status (e.g., disabled and cystic fibrosis). There are also variables that expert clinician input identified as being important to account for, including osteoarthritis, osteoporosis, valvular or rheumatic heart disease and history of opioid use. The full list of variables used in the risk adjustment model can be found in the Measure Codes List, linked at Section S.1.

The 120-day period prior to the start of an episode is used to measure the conditions which most directly impact beneficiaries' health status at the time of the procedure and to capture beneficiaries' comorbidities in the risk adjustment. Additionally, because the relationship between comorbidities' episode cost may be non-linear in some cases (i.e., beneficiaries may also have more than one disease during a hospitalization episode), the model also takes into account a limited set of interactions between HCCs and/or enrollment status variables. The Knee Arthroplasty measure risk adjustment methodology includes only a limited set of interaction terms for two reasons. First, inclusion of too many interaction terms will over-fit the model. Second, the risk-adjustment methodology broadly follows the established CMS-HCC risk-adjustment methodology, which uses similar interaction terms.

Adjustments for Comparability

S.9.1. Inclusion and Exclusion Criteria *Detail initial inclusion/exclusion criteria and data preparation steps (related to clinical exclusions, claim-line or other data quality, data validation, e.g. truncation or removal of low or high dollar claim, exclusion of ESRD patients)*

:

Included populations:

The beneficiary population eligible for the Knee Arthroplasty measure calculation consists of Medicare beneficiaries enrolled in Medicare Parts A and B who received a knee arthroplasty during the performance period, as identified by the episode trigger codes. To be included, the beneficiary must have an episode ending within the performance period to ensure that the beneficiary's claims record contains sufficient fee-for-service data both for measuring spending and for risk adjustment purposes.

Excluded populations:

Episodes are excluded for the following conditions, with the rationale for each provided below:

- The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the 120 days prior to the episode trigger day.
This population is excluded to ensure that we have complete claims data for beneficiaries as there may be other claims (e.g., for services provided under Medicare Part C) that we do not observe in Medicare Parts A and B claims data. Including episodes that do not meet this criterion could potentially misrepresent a clinician's resource use. This exclusion also allows us to accurately construct HCCs for each episode by examining the episode's lookback period without missing claims.
- No attributed clinician is found for the episode.
These episodes are excluded as the measure assesses clinician performance. The measure is intended to assess a homogeneous patient cohort to provide meaningful comparisons between attributed clinicians, so to include these episodes could potentially misrepresent these comparisons.
- The beneficiary's date of birth is missing.

These episodes are excluded as a data cleaning step.

- The beneficiary's death date occurred before the trigger date.

These episodes are excluded as a data cleaning step.

- The beneficiary's death date occurred before the episode ended.

Episodes ending in death are excluded as they are - by definition - truncated episodes and do not have a complete episode window. Including episodes without all observable claims or a complete episode window could potentially make clinicians appear to have lower cost episodes not due to efficiencies of their own performance, but because the data are missing services that would be included in the measure calculation.

- The beneficiary was not enrolled in Medicare Part A and B for the entirety of the 120-day lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. Similarly to above, these episodes are excluded as these beneficiaries may receive services not observed in the data. Including these episode could make the attributed clinician appear to have lower cost episodes due to incomplete data.

- Episodes with inpatient procedures without relevant DRG codes

Episodes will be excluded if the procedure occurred in the inpatient setting and if its concurrent inpatient stay does not have MS-DRG codes that indicate that the reason for admission was for this procedure. These cases are excluded to limit the measure to only capture admissions where the reason for admission is for the knee arthroplasty because cases admitted for other reasons are likely to be more expensive because of the cost of care for the reason for admission as well as for the knee joint replacement.

- Episodes for bilateral partial knee arthroplasties.

Beneficiaries who undergo same-day or staged bilateral partial knee arthroplasties (within 90 days of the first procedure) as identified by modifier codes are excluded from this measure, as rare occurrences. With such a small sample size, these cases are excluded as the effect on expected episode cost is unclear.

- Episodes where the beneficiary has reinsertion/reimplantation of prosthetic knee after infection or spacer during the trigger event or in a 120-day lookback period.

Episodes for procedures for reinsertion or reimplantation either during the trigger event or identified during a 120-day lookback period are excluded because these patients require considerably different clinical care, and are likely to be more expensive.

- Episodes with where the beneficiary has history of infections

Episodes where the beneficiary has recent infections to the knee indicated through CPT/HCPCS and ICD-10 diagnosis codes during a 120-day lookback are excluded because it indicates that these patients require considerably different clinical care, and are likely to be more expensive.

- Episodes classified as outlier cases.

To account for limitations of risk adjustment, episodes predicted to have expected costs that are substantially different from observed costs are excluded as outliers. Specifically, episodes with residuals from the risk adjustment model below the 1st percentile and above the 99th percentile are considered outliers and removed from measure calculation.

S.9.2. Risk Adjustment Type (Select type)

Stratification by risk category/subgroup

If other:

S.9.3. Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets)*

The Knee Arthroplasty measure is stratified into three sub-groups: Partial Knee/Unilateral, Total Knee/Unilateral and Total Knee/Bilateral. These sub-groups represent more homogenous patient cohorts to enable meaningful clinical comparisons based on information available on the trigger claim. These sub-groups are useful in ensuring clinical comparability so that the corresponding cost measure fairly compares clinicians with a similar patient case-mix. A separate risk adjustment model is created for each stratified group, so that clinically meaningful distinctions in the beneficiary population are preserved.

The Knee Arthroplasty measure stratifies cases by partial and total knee replacements because of differing surgical durations and

differing rates of complication, readmission and revision. These differences lead to substantial differences in total costs.

Unilateral partial knee procedures have been outpatient procedures for past years and are likely to be less expensive than unilateral total knee replacements. Bilateral total knee replacements are generally require different resources than unilateral procedures. There will be some services that will not have to be repeated for the second procedure but the patient is likely to require additional rehabilitation, for example at an inpatient rehabilitation hospital stay after the bilateral procedure.

The Knee Arthroplasty measure accounts for the removal of knee arthroplasty procedures from the inpatient-only list through the use of the CPT/HCPCS code to trigger the episode. The current trigger code is based on CPT/HCPCS codes and does not require an inpatient stay. Additionally, risk adjustment for the DRG of the inpatient stay was included, if one is associated with the knee arthroplasty. Specifically, the episode should be included only when the trigger code appears concurrently with MS-DRGs 461, 462, 469 and 470, indicating that the hospital stay was for the knee arthroplasty. With total knee arthroplasties now being allowed in an outpatient setting, patients who receive a knee arthroplasty in an inpatient setting are likely sicker, and clinicians taking care of these patients should not be penalized for the necessary precaution of a longer inpatient stay. The Knee Arthroplasty cost measure also includes Partial Knee Arthroplasty which is sub-grouped in order to create a more homogenous comparison.

S.9.4 Costing method

Detail the costing method including the source of cost information, steps to capture, apply or estimate cost information, and provide rationale for this methodology.

Standardized pricing

The methodology used to payment standardize the Medicare claims used to specify this measure is available for download ("CMS Price (Payment) Standardization") from the URL provided below.

http://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228890990237&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3DDetailed_Mthds_payment-std_041819.pdf&blobcol=urldata&blobtable=MungoBlobs

This direct-download link changes biannually as the documentation is updated; if the link no longer works, the download may also be accessed via the page linked below.

https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228772057350Detailed_Mthds_payment-std_041819.pdf

S.10. Type of score (Select the most relevant):

Ratio

If other:

Attachment:

S.11. Interpretation of Score (Classifies interpretation of a ratio score(s) according to whether higher or lower resource use amounts is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score, etc.)

The Knee Arthroplasty cost measure score is a dollar value that represents a clinician's average payment-standardized risk-adjusted cost to Medicare across all Knee Arthroplasty episodes attributed to them. A value above the national average indicates that on average, the clinician's resource use for this procedure was more expensive than the national average. A value below the national average indicates that on average, the clinician's resource use for this procedure was less expensive than the national average.

We note that this measure – as a cost measure – does not necessarily by itself reflect quality of care. While it does capture consequences of care by including assigned services during the post-trigger period such as for complications, there are other quality metrics that cannot be captured by a cost measure alone. This measure is most meaningful when presented in part of a program such as MIPS where clinicians are also assessed on quality measures.

S.12. Detail Score Estimation (Detail steps to estimate measure score.)

A clinician's Knee Arthroplasty measure score is calculated as the average ratio of observed cost to expected episode cost across a provider's episodes, multiplied by the national average observed episode cost. This calculation is done using episodes from all sub-groups. Further details are provided in Section S.7.2.

Reporting Guidelines

This section is optional and will be available for users of the measure as guidance for implementation and reporting.

S.13.1. Describe discriminating results approach

Detail methods for discriminating differences (reporting with descriptive statistics--e.g., distribution, confidence intervals).

The measure is used in MIPS for the CY 2019 performance period onwards. As such, it has not yet been reported as part of MIPS scoring. However, during measure development, we conducted national field testing where confidential reports containing cost measure performance on the Knee Arthroplasty measure at its draft stage of development (and other episode-based cost measures developed at the same time) were available to clinicians and clinician groups meeting a 10-episode case minimum. The purpose of this field testing was to enable clinicians to become familiar with the measure and to provide feedback on the measure specifications for refinement before CMS considered the measure for use in MIPS. During field testing, a National Summary Data Report was also posted containing summary statistics on the episode-based cost measures, including information on the distribution of TIN and TIN-NPI level measure scores.

S.13.2. Detail attribution approach

Detail the attribution rules used for attributing resources/costs to providers (e.g., a proportion of total measure cost or frequency of visits during the measure's measurement period) and provide rationale for this methodology.

The Knee Arthroplasty episode is attributed to clinicians (TIN-NPIs) billing the episode trigger code. The episode is attributed to a TIN by aggregating all episodes attributed to the NPIs that bill to that TIN. If the same episode is attributed to more than one NPI within a TIN, this episode is only attributed to the TIN once. This allows the measure to be reported to both TINs and TIN-NPIs.

Episodes ending during the performance period are included in a clinician's or clinician group's score. For example, if the performance period is a calendar year, the episode end date (i.e., 90 days after the trigger date) must occur during that calendar year. Requiring episodes to end during the performance period ensures that we have complete claims information for the episode.

S.13.3. Identify and define peer group

Identify the peer group and detail how peer group is identified and provide rationale for this methodology.

Episodes are opened by the presence of trigger codes on Part B physician/supplier claims, so the clinician peer group is limited to those clinicians performing this procedure. This ensures that clinician cost performance for this procedure is being assessed on a homogeneous patient cohort. While this measure was developed for use in MIPS, it can be expanded to other clinician programs.

S.13.4. Sample size

Detail the sample size requirements for reporting measure results.

The Knee Arthroplasty measure will be reported for TINs and TIN-NPIs with 10 or more episodes. The measure is used in the Merit-based Incentive Payment System (MIPS) for MIPS performance period 2019 onwards.

S.13.5. Define benchmarking and comparative estimates

Detail steps to produce benchmarking and comparative estimates and provide rationale for this methodology.

The measure has not been reported yet, as it will be used in the MIPS cost performance category for the 2019 performance period onwards.

Reporting this measure as part of the cost performance category helps to measure clinicians' resource use for knee arthroplasties in the Medicare population, and thereby hold clinicians accountable for their cost effectiveness. There is no reporting/data submission requirement. Combined with measures in the other MIPS performance categories, such as the quality performance category, the Knee Arthroplasty measure allows CMS to assess the value of care and incentivize both achievement and improvement in the provision of high-quality, cost-effective care.

Validity – See attached Measure Testing Submission Form

SA.1. Attach measure testing form

2019_04_16_nqf_testing_form_knee_arthro.docx

Feasibility**F.1. Byproduct of Care Processes**

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

F.1.1. Data Elements Generated as Byproduct of Care Processes.

Generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

F.2. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

F.2.1. To what extent are the specified data elements available electronically in defined fields (i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)

ALL data elements are in defined fields in electronic claims

F.2.1a. 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 other than electronic sources.

F.2.2. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.

Attachment:

F.3. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

F.3.1. Describe what you have learned/modified 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.

Lessons and associated modifications may be categorized into three types: data collection procedures, handling of missing data, and sampling data associated with beneficiaries who died during an episode of care.

Data Collection

Acumen receives claims data directly from the Common Working File (CWF) maintained at the CMS Baltimore Data Center. Medicare claims are submitted by healthcare providers to a Medicare Administrative Contractor (MAC), and are subsequently added to the CWF. However, these claims may be denied or disputed by the MAC, leading to changes to historical CWF data. In rare circumstances, finalizing claims may take many months, or even years. As a result, it is not practical to wait until all claims for a given month are finalized before calculating this measure. Therefore, the time at which a measure developer pulls claims data represents a trade-off between efficiency (accessing the data soon) and accuracy (waiting until most claims are finalized). In order to determine the appropriate "run-out" period for claims data, Acumen has performed testing on the delay between claim service dates and claims data finalization. Based on this analysis, Acumen uses a "run-out" period of three months after the end of the calendar year to collect data for development purposes. MIPS reporting for this cost measure will be done in line with program reporting.

Missing Data

This measure requires complete beneficiary information, and a small number of episodes with missing data are excluded to ensure completeness of data and accurate comparability across episodes (see Section 2b6 of the measure testing form for addition details). For example, episodes where the beneficiary was not enrolled in Medicare Parts A and B for the 120 days prior to the episode start

date are not included in this measure. This enables the risk adjustment model to accurately adjust for the beneficiary's comorbidities using data from the previous 120 days of Medicare claims. Additionally, the risk adjustment model includes a categorical variable for beneficiary age bracket, so episodes for which the beneficiary's date of birth cannot be located are not included in this measure.

Sampling

During measure testing, Acumen noted that episodes in which the beneficiary died prior to the episode end date exhibited different cost distributions to other episodes. In order to avoid this effect impacting clinician scores, this measure does not include episodes for which the beneficiary's date of death occurs prior to the end of the episode window.

F.3.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, and algorithm)?

N/A.

F.3.3. If there are any fees associated with the use of this measure as specified, attach the fee schedule here. (Save file as: F3_3_FeeSchedule)

Usability and Use

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 performance improvement.

U.1.1. Current and Planned Use

| Specific Plan for Use | Current Use (for current use provide URL) |
|-----------------------|--|
| | Payment Program Quality Payment Program Merit-based Incentive Payment System https://qpp.cms.gov/mips/overview |

U.1.2. For each CURRENT use, checked above, provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

Program Name: Quality Payment Program (QPP) Merit-based Incentive Payment System (MIPS)

Sponsor: CMS

Purpose: The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) established the Quality Payment Program. Under the Quality Payment Program, clinicians are incentivized to provide high-quality and high value care through Advanced Alternate Payment Models (APMs) or the Merit-based Incentive Payment System (MIPS). MIPS eligible clinicians will receive a performance-based payment adjustment to their Medicare payment. This payment adjustment is based on a MIPS final score that assesses evidence-based and practice-specific data across the following categories:

1. Quality
2. Improvement activities
3. Advancing care information
4. Cost

As specified in the CY 2019 Physician Fee Schedule final rule (83 FR 59765 through 59776), this measure will be implemented as part of MIPS beginning in the 2019 MIPS performance year and 2021 MIPS payment year.

Geographic Area: U.S.

Number/Percentage of Accountable Entities:

The number of clinicians in the Quality Payment Program varies by performance period. For 2017, there were 1,057,824 MIPS eligible clinicians receiving a MIPS payment adjustment.[1] As clinicians have choices on how to participate in the Quality Payment Program (e.g., through MIPS or the Advanced APMs, as groups or individuals), the exact number and percentage of clinicians who will receive a performance score on this measure will only be confirmed after the end of each performance period.

The number of patients covered by this measure is dependent on whether providers report at the group (TIN) or individual clinician (TIN-NPI) level. The number of patients covered by group reporting is 237,370, while the number of patients covered by individual reporting is 227,070.

[1] CMS, 2017 Quality Payment Program Reporting Experience, <https://qpp-cm-prod-content.s3.amazonaws.com/uploads/491/2017%20QPP%20Experience%20Report.pdf>

Number/Percentage of Accountable Patients: N/A.

U.1.3. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?)

N/A.

U.1.4. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (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.)

N/A.

U.2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation. 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.

DEVELOPMENT: FIELD TESTING

Acumen and CMS conducted a national field test of episode-based cost measures, including the Knee Arthroplasty measure, for a 35-day comment period (October 16 to November 20, 2017). The testing sample for providing field test reports was all clinicians and clinician groups who were attributed 10 or more episodes associated with at least one of eight episode-based cost measures developed during 2017. The measurement period was June 1, 2016, to May 31, 2017. Cost performance information on these episodes were provided in confidential reports available for download on the CMS Enterprise (EIDM) Portal by attributed clinicians and clinician groups.[1]

A sample of 17,557 clinician groups and 48,263 clinicians received a confidential field test report. The testing sample was selected to balance coverage (i.e., a lower minimum number of episodes per attributed clinician increases the number of TIN-NPIs and TINs who would receive reports) and reliability (i.e., a higher minimum number of episodes per clinician or clinician group provides more reliable and meaningful metrics), since a key goal of field testing was to test the measures with as many stakeholders as possible. This sampling technique was used for field testing only and did not determine case minimums used for program implementation.

We provided field test reports for the following number of clinician groups and clinicians. Each report included information for all measures for which the clinician or clinician group was attributed 10 or more episodes.

- Total: 17,557 TINs; 48,263 TIN-NPIs
- Knee Arthroplasty: 3,057 TINs; 10,664 TIN-NPIs

All stakeholders, including those who did not receive a field test report, could review a mock field test report that was posted on the CMS website. Other public documentation posted during field testing included: measure specifications for each measure (comprising a Draft Cost Measure Methodology document and a Draft Measure Codes List file), a Frequently Asked Questions document, and a Fact Sheet.[2] During field testing, Acumen conducted education and outreach activities including National Provider Call webinars, office hours with specialty societies, and Help Desk support.

The purpose of field testing was to provide a voluntary opportunity for clinicians and other stakeholders to provide feedback on: (i)

the draft measure specifications, including each component of the measure (e.g., the clinical validity of assigned services and the trigger codes), (ii) the field test report template (e.g., what information is most meaningful to allow clinicians to make changes to their care practices), and (iii) all accompanying documentation (e.g., the level of detail in specifications documentation). Acumen sought feedback through an online survey, with the option to attach a comment letter in PDF or Word document format.

[1] CMS Enterprise Portal, <https://portal.cms.gov/wps/portal/unauthportal/home/>

[2] These documents were posted to the CMS MACRA Feedback page (<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html>). The field testing fact sheet and FAQs are in a zip file at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2017-field-test-materials.zip>.

IMPLEMENTATION: PRE-RULEMAKING and RULEMAKING

The Knee Arthroplasty measure was implemented in MIPS after going through the pre-rulemaking process and notice-and-comment rulemaking. The measure was submitted to and included in the 2017 Measures Under Consideration (MUC) List. It was then considered by National Quality Forum (NQF)'s Measure Applications Partnership (MAP) Clinician Workgroup and Coordinating Committee in December 2017 and January 2018, respectively. The final recommendation from the MAP was 'conditional support for rulemaking,' with the condition of NQF endorsement.

The measure was proposed for use in the MIPS cost performance category in the CY 2019 Physician Fee Schedule proposed rule.[1] Measure specifications were publicly posted and linked to from the proposed rule. A National Summary Data Report containing information about the measure performance (e.g., measure score distributions by different provider characteristics) was also publicly posted. Stakeholders submitted comments on the proposed rule during a 60-day public comment period. CMS considered these comments and finalized the measure for use in MIPS from the CY 2019 performance period onwards in the CY 2019 Physician Fee Schedule final rule.[2]

[1] The CY 2019 Physician Fee Schedule proposed rule can be found here:

<https://www.federalregister.gov/documents/2018/07/27/2018-14985/medicare-program-revisions-to-payment-policies-under-the-physician-fee-schedule-and-other-revisions>

[2] The CY 2019 Physician Fee Schedule final rule can be found here:

<https://www.federalregister.gov/documents/2018/11/23/2018-24170/medicare-program-revisions-to-payment-policies-under-the-physician-fee-schedule-and-other-revisions>

U.2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

FIELD TESTING:

National field testing was organized for the purpose of gathering targeted comments on the Knee Arthroplasty measure. During the feedback period, field test reports were accessed by accounts corresponding to a total of 1,364 clinician groups (TINs) and 10,628 clinicians (TIN-NPIs). After field testing, the comments received on the measure were summarized for the Clinical Subcommittee to consider in making refinements. Field test reports continued to be available on the CMS Enterprise Portal until September 2018.

The following sections offer more details on the contents of each report and describe the education and outreach efforts associated with the field testing feedback period.

Data Provided During Field Testing

Each field test report Excel file contained the following sheets, which were described in more detail in Appendix C of the field test reports ("How to Interpret this Report"):

- Summary
 - o High-level information on the performance of the TIN or TIN-NPI across all episodes within each measure attributed to said TIN or TIN-NPI
 - o Metrics listed in this tab related to the cost measure score:
 - ? Episode count
 - ? Average episode risk score percentile
 - ? Cost measure score (average risk-adjusted cost to Medicare for that measure)
 - ? National average cost measure score and percent difference between the TIN/TIN-NPI's score and the national average
- Results for Each Measure

- o Understanding Your Cost Measure Score: information from the Summary tab in context
- o Breakdown of Cost Measure Score by Episode Sub-Group: comparison of the TIN/TIN-NPI's average risk-adjusted cost to Medicare to the national average risk-adjusted cost to Medicare for the measure as a whole and separately for each episode sub-group
- ? Episode sub-groups are divisions within a cost measure's episode group that define more homogenous patient cohorts to ensure clinical comparability (i.e., the cost measure fairly compares like patients)
- o Breakdown of Episodes by Episode Sub-Group for Your TIN/TIN-NPI and National Average: comparison of the allocation of the TIN/TIN-NPI's episodes to the various sub-groups within the overall episode group to the average allocation across episodes for TINs/TIN-NPIs nationally
- o Breakdown of Part B Physician/Supplier Episode Cost by Your TIN/TIN-NPI vs. Other TINs/TIN-NPIs: average share of episode costs that came from the evaluated TIN/TIN-NPI versus other TINs/TIN-NPIs and average of each share across episodes for TINs/TIN-NPIs nationally
- o Breakdown of Utilization and Cost by Selected Clinical Theme: TIN's/TIN-NPI's service utilization and costs by "clinical themes" (clinical categorizations of the services assigned to episode costs during the episode window)[1]
 - Appendix A for Each Measure
- o More detailed information on potential cost drivers in the TIN/TIN-NPI's episodes
- o Breakdown of Utilization and Cost by Medicare Setting and Service Category: analysis of utilization and cost for the measure, both for all services and by specific service categories[2]
- o Breakdown of Utilization and Cost for Physician/Supplier Part B Claims: same comparison of utilization and cost as given in "Breakdown of Utilization and Cost by Medicare Setting and Service Category" above (i.e., (i) the national average, (ii) TINs/TIN-NPIs in the same risk bracket, and (iii) the evaluated TIN/TIN-NPI), but by top 5 most billed services and by risk bracket
- o Breakdown of Utilization and Cost for Inpatient Claims: same information as in "Breakdown of Utilization and Cost for Physician/Supplier Part B Claims" for inpatient claims assigned to the TIN/TIN-NPI's episode costs
 - Appendix B
- o Detailed episode-level information for all episodes attributed to the TIN/TIN-NPI across all measures in the report
- o Data across six major categories: (i) Episode Costs, (ii) Beneficiary Information, (iii) Attributed Clinician(s), (iv) Evaluation and Management Visits Performed During Episode, (v) Physician Fee Schedule Costs to Medicare Billed During Episode, and (vi) Other Providers Rendering Care Within the Episode

[1] Definitions of the clinical themes are available in the "SA_" tabs of the Measure Codes List file for the measure, downloadable from the QPP Resource Library at this link: <https://qpp-cm-prod-content.s3.amazonaws.com/uploads/344/2019%20Cost%20Measure%20Code%20Lists.zip>

[2] Definitions of the various categories of services presented in this table can be found on page 438, Table C.2 of the "Detailed Methods of the 2015 Supplemental Quality and Resource Use Reports (QRURs)" document available here: <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2015-SQRUR-Detailed-Methods.pdf>

Education and Outreach

Acumen directly conducted outreach via email to tens of thousands of stakeholders using the stakeholder contact list developed through previous education and outreach and Clinical Subcommittee recruitment efforts, as well as CMS, QPP, and other available listservs. Outreach emails included:

- Targeted messages to a small number of specialty societies whose members we anticipated would be attributed a report, to assist in reaching their members about field testing
- Targeted emails to available contact details linked to a TIN or TIN-NPI that received a field test report
- General emails to contacts from clinician and healthcare provider organizations, noting that we sought feedback from all stakeholders even if they did not receive a confidential report
- General emails to all our contacts in clinician and healthcare provider organizations to inform them about the opportunity to join National Provider Calls (NPCs)

Acumen and CMS hosted two office hours sessions on September 14 and 18, 2017, to provide an overview of field testing to specialty societies, discuss what information their members would be particularly interested in, and answer any questions. After the webinars, Acumen also prepared and distributed measure summaries that societies could use to inform their members of the basic specifications on which Acumen was requesting input. Between the two webinars, there were 31 attendees affiliated with at least 21 specialty societies, including representatives from the American Academy of Orthopaedic Surgeons, the American Academy of Physical Medicine and Rehabilitation, and the American Medical Association, among others.

During the field testing feedback period, Acumen organized an inquiry management strategy with the Physician Value and QPP Help Desks; Acumen directly handled more than 160 inquiries during the feedback period.

Acumen and CMS hosted two National Provider Calls on October 30, 2017, and November 2, 2017, to engage clinicians and other stakeholders during field testing. The two webinars, both covering the same content, consisted of an hour-long presentation, outlining (i) the cost measure development activities, (ii) how to access the confidential field test reports, and (iii) the contents of the reports. The presentation was followed by a 30-minute Q&A session. In total, approximately 1,000 people attended one of the webinars and around 120 comments and questions were received via webinar chat and on the phone.

PRE-RULEMAKING:

There was a public comment period after the release of the Measures Under Consideration (MUC) list from November 30, 2017 to December 7, 2017, prior to the MAP Clinician Workgroup meeting. The MAP Clinician Workgroup met on December 12, 2017, to consider measure specifications and testing updates. In accordance with MAP procedure, these documents were not publicly released but were made available to MAP members. Following the release of the Clinician Workgroup's preliminary recommendation, the report was open for a public comment period from December 21, 2017, to January 11, 2018. The MAP Coordinating Committee met on January 25-26, 2018, to consider these comments alongside the Clinician Workgroup's recommendation. Both MAP meetings were open to the public.

RULEMAKING:

During the public comment period for the proposed rule from July 12, 2018, to September 10, 2018, stakeholders could review the proposed rule language, measure specifications, and National Summary Data Report when submitting comments. CMS conducted email outreach via its listserv to notify stakeholders about the release of the proposed rule.

U.2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1. Describe how feedback was obtained.

FIELD TESTING:

In total, Acumen received 219 survey responses and 53 comment letters, including many from specialty societies representing large numbers of potentially attributed clinicians.

Survey responses and comment letters were collected via an online survey, which proved advantageous in reaching a wider audience, increasing the amount and variety of feedback provided, and facilitating a faster turnaround for the measure development team to process and operationalize feedback. The survey was divided into four sections for general and detailed questions on the reports themselves, questions on the supplemental documentation, and questions on the measure specifications. Questions in the survey included Likert scales specific to the report, process, and measure components; multiple-choice questions; and open response questions. The survey was designed to take 20-30 minutes, but allowed flexibility based on a stakeholder's use of open-ended responses and the number of measures on which they chose to provide feedback.

Inquiries were also registered and feedback submitted via email to macra-episode-based-cost-measures-info@acumenllc.com, Physician Value and QPP Help Desks, and verbally and via webinar chat at the NPCs and office hours.

PRE-RULEMAKING:

CMS received over 40 comments on the eight episode-based cost measures included in the 2017 Measures Under Consideration List. This included six comments for the Knee Arthroplasty Cost Measure. After the MAP Clinician Workgroup meeting in December 2017, there was another public comment period on their preliminary recommendations, which received over 20 comments across the eight measures, with four comments specific to the Knee Arthroplasty Cost Measure. These public comment periods were facilitated by NQF. Stakeholders were able to submit their comments via the NQF website.

RULEMAKING:

CMS received over 15,368 comments on the CY 2019 Physician Fee Schedule proposed rule. A search on the [regulations.gov](https://www.regulations.gov) website returns 242 results for "episode-based cost measure" as a rough approximation of the number of comments on the eight episode-based cost measures during rulemaking. Stakeholders could submit comments through the Federal Register website or via mail.

U.2.2.2. Summarize the feedback obtained from those being measured.

FIELD TESTING:

The publicly available Field Testing Feedback Summary Report[1] presents all feedback gathered during the field testing period. The following list synthesizes some of the key points that were raised through the field testing feedback period:

- Stakeholder engagement and involvement is an important aspect of the measure development process. Stakeholders expressed appreciation for the opportunity to provide feedback during field testing and for CMS' continued effort to involve stakeholders in the measure development process, such as convening Clinical Subcommittees to seek an extensive amount of clinical input in constructing these measures. Commenters urged CMS to continue to work closely with specialty societies and other involved stakeholders.
- Provide additional time for stakeholders to review materials and provide feedback during field testing. According to some stakeholders, the October to November 2017 field testing feedback period was too short given the large amount of new information that was presented and suggested that the period be extended or be kept open.
- Accessing the confidential field test reports from the CMS Enterprise Portal presented many challenges. Some stakeholders noted that they faced difficulties creating accounts and downloading their confidential field test reports from the portal that may have had a negative impact on the number of clinicians who took part in field testing.
- While some stakeholders believed the field test report presented useful information for understanding clinician cost measure performance, they also highlighted areas for improvement in regard to providing actionable information. Stakeholders praised the navigability and the inclusion of useful information in the report. However, some stakeholders also expressed concerns with the comprehensibility of the report and its usefulness in terms of providing actionable information for clinicians.
- Stakeholder feedback received on the supplemental field testing materials was mixed, with some stakeholders finding them helpful and informative and others believing the materials were too complex. Some stakeholders found the supplemental field testing materials informative, providing helpful information on field testing and the specifications of the cost measures. Some stakeholders believed that the materials were not detailed enough. However, many noted that the materials were comprehensive but too lengthy and complex, and they believed the amount of information was overwhelming to absorb within the field testing feedback period.

The aforementioned report additionally contains measure-specific feedback, which was used as the basis for the post-field testing measure refinements discussed in U.2.3. At a high level, feedback included the following recommendations:

- Refinements to trigger codes, attribution, sub-groups, episode windows, assigned services, risk adjustment variables, exclusions, and alignment of cost with quality
- Adding/removing certain trigger codes and assigned services, further sub-grouping, and revising the attribution methodology

Stakeholders also noted that the level of clinician engagement in the development of these episode-based cost measures is a significant improvement over the development process for earlier cost measures.

Feedback collected via email, Help Desk, and at the NPCs and office hours covered a wide range of topics, such as:

- Email and Help Desk inquiries: accessing field test reports, MIPS cost performance category, cost measures for chronic conditions, interpreting field test reports, using the online survey, payment standardization
- NPC and office hours comments and questions: risk adjustment methodology, supplemental field test resources, field test methodology, quality alignment, future cost measures

[1] The report can be downloaded at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2018-field-testing-feedback-summary-report.pdf>

PRE-RULEMAKING:

The MAP gives feedback on performance measures from a wide variety of perspectives, with representatives including “consumers, businesses and purchasers, laborers, health plans, clinicians and providers, communities and states, and suppliers.”[2] The Clinician Workgroup specifically aims to “ensur[e] the alignment of measures and data sources to reduce duplication and burden, identif[y] the characteristics of an ideal measure set to promote common goals across programs, and implemen[t] standardized data elements.”[3]

[2] National Quality Forum, Measure Applications Partnership

https://www.qualityforum.org/Setting_Priorities/Partnership/Measure_Applications_Partnership.aspx

[3] National Quality Forum, MAP Member Guidebook

<http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=80515>

RULEMAKING/PUBLIC COMMENT:

CMS received comments on the proposed episode-based cost measures during the public comment period for the CY 2019 Physician Fee Schedule proposed rule. There was support from several commenters on the proposed adoption of the episode-based cost measures in MIPS. Commenters provided feedback on the development process, including voicing support for the development of episode-based cost measures through a transparent process that engages with stakeholders and submitting critiques of the short timeline clinicians are given to understand and gain experience with the measures before they are used in the program. CMS also received comments supporting the submission of the episode-based cost measures for NQF endorsement prior to their use in the program. Measure-specific comments were also received on the specifications of the measures, which CMS and Acumen reviewed to determine whether changes needed to be made to the specifications of the measures. For more detailed information on the comments received on the measures as part of the proposed rule public comment period, please see the episode-based cost measures section in the CY 2019 Physician Fee Schedule final rule for a summary of the public comments received along with CMS responses: <https://www.federalregister.gov/d/2018-24170/p-2965>.

U.2.2.3. Summarize the feedback obtained from other users.

PRE-RULEMAKING:

The MAP recognized the importance of this Knee Arthroplasty cost measure and conditionally supported the measure pending NQF endorsement. During NQF endorsement review, the MAP encouraged the Cost and Resource Use Standing Committee to specifically consider the appropriateness of the risk adjustment model to ensure clinical and social risk factors are reviewed and included when appropriate. MAP cautioned about the potential stinting of care and noted that appropriate risk adjustment could help safe guard against this practice. Additionally, MAP expressed concern over the precision of the cohort definition and whether there was a sufficiently large cost performance distribution in this measure.

U.2.3. Describe how the feedback described in 4a2.2 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not

FIELD TESTING:

Careful consideration was given to all feedback gathered during field testing, and several updates were made to the measure based on the recommendations of field test commenters and a Clinical Subcommittee comprised of subject matter and measure-development experts.

After completing field testing, the feedback provided through the survey and comment letters was compiled into a measure-specific report, which was then provided to the Clinical Subcommittee (CS) that provided the bulk of measure development input. CS members then discussed and voted on which of the proposed specifications updates should be implemented in the finalized measure. More specifically, this process included the following steps for each of two webinars:

- Pre-webinar production of summary sheets, first of applicable field testing feedback and then of first-round measure refinements
- The webinar itself, to gather first substantive and then non-substantive measure refinement feedback
 - o CS members discussed each update suggested by field testing commenters to determine whether, based on their best clinical input, they should recommend implementation of the change or not.
 - o In some cases, CS members acknowledged the validity of the suggestion, but felt they had already addressed the commenter(s) concerns.
- A survey to gather CS member input
 - o For the purposes of considering which measure specifications changes to implement, CS consensus was defined as >60% agreement.
- Incorporation of CS input into final measure specifications

The changes to the Knee Arthroplasty measure made as a result of field testing feedback are as follows:

- Exclusions:
 - o Add exclusion for revisions and spacers
- Service Assignment:
 - o Do not assign costs related to the following pre-operative services: cardiac cath, cardiac stress test, nuclear stress test, pulmonary testing
 - o Change post-trigger window for assigning costs related to UTI from <30 days to <15 days
 - o Assign costs for all HCPCS codes within CCS 213 and 215

RULEMAKING/PUBLIC COMMENT:

During the public comment period for the CY 2019 Physician Fee Schedule proposed rule, stakeholders submitted comments on the

proposed episode-based cost measures, including on the Knee Arthroplasty measure. While we received feedback on the proposed measures generally, as described in Section U.2.2.2, there was no measure-specific feedback received on the specifications of this measure. Therefore, the measure was finalized as proposed.

U.3.1. Progress on Improvement. (Not required for initial endorsement unless available.) Performance results on this measure (current and over time) should be provided in IM.1.2 and IM.1.4.

Discuss:

- Purpose Progress (trends in performance results)
- Geographic area and number and percentage of accountable entities and patients included

N/A.

U.3.2. If no improvement was demonstrated, what are the reasons? 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.

N/A.

U.4.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

While the measure has technically been implemented into the MIPS program, the measure results are first scheduled to be calculated for performance year 2019 (payment year 2021), and thus no unexpected consequences can be identified at this time.

U.4.2. Please explain any unexpected benefits from implementation of this measure.

While the measure has technically been implemented into the MIPS program, the measure results are first scheduled to be calculated for performance year 2019 (payment year 2021), and thus no unexpected consequences can be identified at this time.

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.

H.1. Relation to Other NQF-endorsed Measures

If there are related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

H.1.1. List of related or competing measures (selected from NQF-endorsed measures)

1609 : ETG Based HIP/KNEE REPLACEMENT cost of care measure

H.1.2. If related or competing measures are not NQF endorsed please indicate measure title and steward.

Related NQF-Endorsed Measures:

There are no NQF-endorsed cost measures with the same focus or the same target population.

Competing NQF-Endorsed Measures:

There are currently no NQF-endorsed measures that address both this same measure focus AND this same target population.

Related Non-NQF-Endorsed Measures:

- Title: Hospital-level, risk-standardized payment associated with a 90-day episode of care for elective primary total hip and/or total knee arthroplasty (THA/TKA)
 - o Measure Steward: CMS
 - o NQF #3474 (currently undergoing review, recommended for endorsement by the NQF Cost and Efficiency Standing Committee)
- Title: ETG Based HIP/KNEE REPLACEMENT cost of care measure
 - o Measure Steward: Optum
 - o NQF #1609 (endorsement removed)

Competing Non-NQF-Endorsed Measures:

| |
|---|
| There are currently no non-NQF-endorsed measures that address both this same measure focus AND this same target population. |
| H.2. Harmonization H.2.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized? H.2.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden. |
| H.3. Competing Measure(s) H.3.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.) N/A. There are currently no measures that have both the same focus and target population. This Knee Arthroplasty measure evaluates clinicians' and clinician groups' risk-adjusted episode cost. The target population is Medicare beneficiaries enrolled in Medicare fee-for-service and who receive an elective knee arthroplasty that triggers a Knee Arthroplasty episode. The cohort for this cost measure is also further refined by the definition of the episode group and measure-specific exclusions. |

| |
|---|
| Contact Information |
| Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services Co.2 Point of Contact: Joel, Andress, joel.andress@cms.hhs.gov, 410-786-5237- Co.3 Measure Developer if different from Measure Steward: Acumen, LLC Co.4 Point of Contact: Bingle, Luo, ccsq-macra-support@acumenllc.com, 650-558-8882- |
| Additional Information |
| Ad.1 Workgroup/Expert Panel involved in measure development List the workgroup/panel members' names and organizations. Describe the members' role in measure development. Acumen convened multiple stakeholder and expert groups to contribute to the measure development process, including Clinical Subcommittees and a Technical Expert Panel (TEP). Clinical Subcommittees convened between May 2017 and January 2018 made recommendations on all components of the episode-based cost measures, including what diagnoses and/or procedures should trigger and define an episode, which services should be assigned to an episode, what patient populations should be excluded, and which clinical characteristics should be accounted for in the risk adjustment model. The TEP, which met four times between August 2016 and August 2017, served a high-level advisory role and provided cross-measure guidance on the overall direction of measure development. Technical Expert Panel Members: Adolph Yates, American Academy of Orthopaedic Surgeons Alan Lazaroff, American Geriatrics Society Allison Madson, American Society of Cataract and Refractive Surgery Alvia Siddiqi, American Academy of Family Physicians Anupam Jena, Harvard Medical School Caroll Koscheski, American College of Gastroenterology Chandy Ellimoottil, American Urological Association Diane Padden, American Association of Nurse Practitioners Dyane Tower, American Podiatric Medical Association Edison A. Machado, Jr., The American Health Quality Association |

Jackson Williams, Dialysis Patient Citizens
 James Naessens, Mayo Clinic
 John Bulger, American Osteopathic Association
 Juan Quintana, American Association of Nurse Anesthetists
 Kata Kertesz, Center for Medicare Advocacy
 Kathleen Blake, American Medical Association
 Mary Fran Tracy, National Association of Clinical Nurse Specialists
 Parag Parekh, American Society of Cataract and Refractive Surgery
 Patrick Coll, University of Connecticut Health Center
 Shelly Nash, Adventist Health System
 Sophie Shen, Johnson and Johnson Health Care Systems, Inc.

Musculoskeletal Disease Management - Non-Spine Clinical Subcommittee Members:

Adam Rana, American Association of Hip and Knee Surgeons
 Adolph Yates, American Academy of Orthopaedic Surgeons
 Alex Limanni, American College of Rheumatology
 Ammar Sarwar, Society of Interventional Radiology
 Andrew Gordon, American Academy of Physical Medicine and Rehabilitation
 Bela Pandit, The American College of Foot and Ankle Surgeons
 Daniel Moon, American Orthopaedic Foot and Ankle Society
 Daniel Wessell, American College of Radiology
 David Friedenson, American College of Emergency Physicians
 Dennis Rivenburgh, American Academy of Physician Assistants
 Dheeraj Mahajan, AMDA - The Society for Post-Acute and Long-Term Care Medicine
 Edward Mariano, American Society of Anesthesiologists
 Heather Smith, American Physical Therapy Association
 Jennifer Koontz, American Academy of Family Physicians
 Jeremy Furniss, American Occupational Therapy Association
 Juan Quintana, American Association of Nurse Anesthetists
 Kenneth Hunt, American Orthopaedic Foot and Ankle Society
 Kevin Klauer, American College of Emergency Physicians
 Kirk Whetstone, American Academy of Physical Medicine and Rehabilitation
 Marc DeHart, American Academy of Orthopaedic Surgeons
 Mark Levine, American Geriatrics Society
 Michael DePalma, American Academy of Physician Assistants
 Michael Wasserman, American Geriatrics Society
 Michael Zychowicz, American Academy of Nurse Practitioners
 Robert (Dale) Blasier, North American Spine Society
 Robin Kamal, American Academy of Orthopaedic Surgeons
 Steven Schmitt, Infectious Diseases Society of America
 William Beach, American Orthopaedic Society for Sports Medicine

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

Ad.3 Month and Year of most recent revision:

Ad.4 What is your frequency for review/update of this measure?

Ad.5 When is the next scheduled review/update for this measure?

Ad.6 Copyright statement:

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: