

Renal or Ureteral Stone Surgical Treatment Measure

Measure Justification Form

June 2019



Contents

1.0	Introduction	4
1.1	Project Title and Overview	4
1.2	Measure Name	4
1.3	Type of Measure	4
2.0	Importance.....	5
2.1	Evidence to Support the Measure Focus	5
2.1.1	Measure Description	5
2.1.2	Evidence for Measure Focus	5
2.2	Performance Gap	6
2.2.1	Rationale	6
2.2.2	Performance Scores.....	6
3.0	Scientific Acceptability	8
3.1	Data Sample Description	8
3.1.1	Type of Data Used for Testing.....	8
3.1.2	Specific Dataset Used for Testing	8
3.1.3	Dates of the Data Used in Testing.....	8
3.1.4	Levels of Analysis Tested.....	8
3.1.5	Entities Included in the Testing and Analysis.....	8
3.1.6	Patient Cohort Included in the Testing and Analysis.....	8
3.1.7	Sample Differences	10
3.1.8	Social Risk Factors Included in Analysis	10
3.2	Reliability Testing.....	10
3.2.1	Level of Reliability Testing	10
3.2.2	Method of Reliability Testing	10
3.2.3	Statistical Results from Reliability Testing	11
3.2.4	Interpretation	12
3.3	Validity Testing	12
3.3.1	Level of Validity Testing.....	12
3.3.2	Method of Validity Testing	12
3.3.3	Statistical Results from Validity Testing	14
3.3.4	Interpretation	15
3.4	Exclusions Analysis	15
3.4.1	Method of Testing Exclusions.....	15
3.4.2	Statistical Results from Testing Exclusions.....	16
3.4.3	Interpretation	16
3.5	Risk Adjustment or Stratification	17
3.5.1	Method of Controlling for Differences	17
3.5.2	Conceptual, Clinical, and Statistical Methods	19
3.5.3	Conceptual Model of Impact of Social Risks.....	19
3.5.4	Statistical Results	19
3.5.5	Analyses and Interpretation in Selection of Social Risk Factors.....	20
3.5.6	Method for Statistical Model or Stratification Development.....	21
3.5.7	Statistical Risk Model Discrimination Statistics	21
3.5.8	Statistical Risk Model Calibration Statistics	22
3.5.9	Statistical Risk Model Calibration – Risk Decile.....	22
3.5.10	Results of Risk Stratification Analysis.....	22
3.5.11	Interpretation.....	23
3.6	Identification of Meaningful Differences in Performance.....	23

3.6.1 Method	23
3.6.2 Statistical Results	23
3.6.3 Interpretation	24
3.7 Missing Data Analysis and Minimizing Bias.....	24
3.7.1 Method	24
3.7.2 Missing Data Analysis	24
3.7.3 Interpretation	25
4.0 Feasibility.....	26
4.1 Data Elements Generated as Byproduct of Care Processes	26
4.2 Electronic Sources	26
4.3 Data Collection Strategy	26
4.3.1 Data Collection Strategy Difficulties.....	26
5.0 Usability and Use	27
5.1 Use	27
5.1.1 Current and Planned Use	27
5.1.2 Feedback on the Measure and Development Process	27
5.2 Usability	31
5.2.1 Improvement	31
5.2.2 Unexpected Findings.....	31
5.2.3 Unexpected Benefits	31
6.0 Related and Competing Measures	32
6.1 Relation to Other Cost Measures	32
6.2 Harmonization.....	32
6.3 Competing Measures	32
Contact Information	33

1.0 Introduction

This Measure Justification Form (MJF) provides results for the testing and evaluation of the Renal or Ureteral Stone Surgical Treatment measure. The MJF is intended to provide detailed information about the testing conducted on this measure, and accompanies the Measure Methodology and Measure Codes List files, which together comprise the specifications for this cost measure.¹

1.1 Project Title and Overview

The Centers for Medicare & Medicaid Services (CMS) has contracted with Acumen, LLC to develop care episode and patient condition groups for use in cost measures to meet the requirements of the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). The contract name is “MACRA Episode Groups and Cost Measures.” The contract number is HHSM-500-2013-13002I, Task Order HHSM-500-T0002.

1.2 Measure Name

Renal or Ureteral Stone Surgical Treatment Episode-Based Cost Measure

1.3 Type of Measure

Cost/Resource Use

¹ CMS, “Renal or Ureteral Stone Surgical Treatment Measure Methodology,” *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2019-revised-ebcm-measure-specs.zip>.

CMS, “Renal or Ureteral Stone Surgical Treatment Measure Codes List,” *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2019-revised-ebcm-measure-specs.zip>.

2.0 Importance

2.1 Evidence to Support the Measure Focus

2.1.1 Measure Description

The Renal or Ureteral Stone Surgical Treatment cost measure evaluates clinicians' risk-adjusted cost to Medicare for beneficiaries who receive surgical treatment for renal or ureteral stones. The cost measure score is a clinician's average risk-adjusted cost for the episode group 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 90 days prior to the clinical event that opens or 'triggers' the episode through 30 days after the trigger. Beneficiary populations eligible for the Renal or Ureteral Stone Surgical Treatment measure include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

2.1.2 Evidence for Measure Focus

Policymakers contend that an estimated 80 percent of overall health care costs are attributable to decisions made by clinicians.² 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 with opportunities for improvement. These include improving the quality of care in outpatient settings to mitigate costs and establishing treatment guidelines to reduce procedure variation and recurrence.

Shifting stone treatment to outpatient settings and ensuring that outpatient treatment is comparable to the quality of care received in the inpatient setting is one method to improve both cost-efficiency and outcomes. Advanced technology has improved the efficiency of stone surgery, allowing more procedures to be done in the outpatient setting.³ As a result, the number of ambulatory evaluation and management visits for urinary stones has increased while the rate of inpatient hospitalizations has decreased. In 2013, approximately 23,000 Medicare beneficiaries were hospitalized with a primary diagnosis of kidney stones while approximately 1.1 million received ambulatory and outpatient evaluation and management, a 75 percent increase from 2004.⁴ With this shift, optimizing the quality of care in outpatient settings could help preserve or increase the cost advantages of outpatient care.

One way to improve the quality of care in outpatient settings is to improve adherence to medical guidelines, which, if not followed, could delay the provision of urgent interventions and lead to additional costs. A study examining adherence to clinical guidelines found that guideline-recommended care was absent or varied widely among patients who received outpatient

² 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.

³ Hollingsworth, John M, Zaojun Ye, et al., "Urologists Ownership of Ambulatory Surgery Centers and Urinary Stone Surgery Use." Health Services Research Journal vol. 44, no. 4, 2009, pp. 1370-1384.

⁴ "Urologic Diseases in America. Kidney Stones." National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2017.

services for kidney stone treatment.⁵ Only 40 percent of emergency department visits completed all three guideline-based laboratory tests, with utilization of each test widely varying, and pharmacologic therapy for facilitating stone passage was prescribed for only 17 percent of eligible visits. These shortcomings to care delivery could increase costs and temporary disability, suggesting there is an opportunity for improvement and substantial cost savings.⁶

2.2 Performance Gap

2.2.1 Rationale

In 2013, 23,000 Medicare beneficiaries were hospitalized with a primary diagnosis of kidney stones, and approximately 1.1 million beneficiaries with a primary diagnosis of kidney stones received ambulatory and outpatient (OP) evaluation and management care.⁷ It is estimated that the total expenditure among Medicare beneficiaries 65 and older for treatment of urinary tract stones exceeds \$1 billion each year.⁸ The Renal or Ureteral Stone Surgical Treatment episode-based cost measure was recommended for development by an expert clinician committee—the Urologic Disease Management 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. Based on the initial recommendations from the Clinical Subcommittee, the subsequent measure-specific workgroup provided extensive, detailed input on this measure.

2.2.2 Performance Scores

Performance scores are provided for 1,661 clinician group practices (identified by Tax Identification Number [TIN]) and 4,158 practitioners (identified by combination of TIN and National Provider Identifier [NPI]). These counts represent attributed clinicians and clinician groups billing Part B Physician/Supplier claims under a Merit-based Incentive Payment System (MIPS) eligible clinician specialty, and do not reflect other MIPS eligibility criteria (e.g., Advanced Alternative Payment Model participation). This table uses a testing volume threshold of 10 episodes.

⁵ Scales, Jr. Charles D, Jonathan Bergman, et al., “Quality of Acute Care for Patients with Urinary Stones in the United States.” *Urology* vol. 86, no. 5, 2015, pp. 914-921.

⁶ Strobe, Seth A, J Stuart Wolf Jr. et al., “Changing Practice Locations for Upper Urinary Tract Stone Disease.” *The Journal of Urology*, vol. 182, no. 3, 2009, pp. 1005-1011.

⁷ “Urologic Diseases in America. Kidney Stones.” National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2017.

⁸ Table 14-46. Economic Impact of Urologic Disease. In: Chapter 14. Litwin MS, Saigal CS, editors. *Urologic Diseases in America*. US Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases. Washington, DC: US Government Printing Office, 2012; NIH Publication No. 12-7865 pp. 486.

Table 1: Distribution of Performance Scores

Metric	TIN	TIN-NPI
Mean score	\$6,192	\$6,170
Standard deviation	\$758	\$796
Score IQR	\$814	\$959
Score percentile		
10 th	\$5,422	\$5,272
20 th	\$5,650	\$5,525
30 th	\$5,809	\$5,738
40 th	\$5,945	\$5,904
50 th	\$6,082	\$6,074
60 th	\$6,235	\$6,262
70 th	\$6,434	\$6,474
80 th	\$6,677	\$6,745
90 th	\$7,123	\$7,166

3.0 Scientific Acceptability

3.1 Data Sample Description

3.1.1 Type of Data Used for Testing

Medicare administrative claims, Long-Term Minimum data set (MDS), enrollment database (EDB), and Common Medicare Environment (CME)

3.1.2 Specific Dataset Used for Testing

The Renal or Ureteral Stone Surgical Treatment measure uses Medicare Part A and Part B claims data 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 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 based on the data from the MDS. Specifically, 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 CME are used in analyses evaluating social risk factors in risk adjustment.

3.1.3 Dates of the Data Used in Testing

The measurement period includes Renal or Ureteral Stone Surgical Treatment episodes ending from January 1, 2017 through December 31, 2017

3.1.4 Levels of Analysis Tested

Individual clinician (identified by combination of TIN and NPI) and clinician group/practice (identified by TIN).

3.1.5 Entities Included in the Testing and Analysis

1,661 clinician group practices and 4,158 practitioners were included in the analyses. Clinicians and clinician groups were included in testing if they were attributed 10 or more Renal or Ureteral Stone Surgical Treatment episodes during the measurement period. Episodes from all 50 States and D.C. in the following settings were included: acute inpatient (IP) hospitals, hospital outpatient departments (HOPD), ambulatory/office-based care centers, and ambulatory surgical centers (ASC).

3.1.6 Patient Cohort Included in the Testing and Analysis

83,307 Medicare beneficiaries (from 99,613 episodes) were included in TIN level testing and analysis, and 71,030 beneficiaries (from 85,207 episodes) were included in TIN-NPI level measure testing.

The beneficiary population eligible for the Renal or Ureteral Stone Surgical Treatment measure calculation consists of Medicare beneficiaries enrolled in Medicare Parts A and B (but not Part C) who received surgical treatment for renal or ureteral stones during the measurement period, as identified by the episode trigger Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) codes on Part B Physician/Supplier claims. Beneficiaries and their episodes were included in the sample if they met a set of inclusion criteria (listed below) meant to ensure completeness of data and to focus the measure on a clinically homogeneous cohort of patients receiving renal or ureteral stone surgical treatment.

The inclusion criteria are:

- The beneficiary has Medicare as their primary payer for the entire episode window, as well as the 120 days prior to the trigger day (the 120-day lookback period).
- The beneficiary was continuously enrolled in Medicare Parts A and B, and not enrolled in Part C, for the entirety of the episode window and the 120-day lookback period.
- The beneficiary date of birth is not missing.
- The beneficiary death date did not occur before episode end.
- The beneficiary has a sufficient 120-day lookback period.
- The episode can be attributed to at least one main clinician.
- The episode does not contain two types of trigger code (i.e., the episode does not have any of the following combinations occurring on the trigger date: (i) Extracorporeal Shock Wave Lithotripsy with Ureteroscopy, (ii) Percutaneous Nephrostolithotomy with Extracorporeal Shock Wave Lithotripsy, and (iii) Percutaneous Nephrostolithotomy with Ureteroscopy))
- The episode trigger claim occurred in an IP, OP, office, or ASC setting based on its place of service.
- If the episode trigger has a concurrent IP stay, that stay occurs in a short-term stay acute hospital as defined by subsection (d).⁹
- If the trigger event was performed inpatient, the IP stay was billed as a transurethral procedure.
- The beneficiary did not have a diagnosis of calculus of lower urinary tract on or before the trigger date.
- The beneficiary underwent no more than one OP non-endoscopic urologic procedure on the trigger day.
- The episode is not an outlier case.

To determine whether the Renal or Ureteral Stone Surgical Treatment measure's inclusion criteria distort patient characteristics on episodes, we produced and analyzed distributions of patient characteristics (age, race, sex, dual eligibility status, income, unemployment, hierarchical condition categories [HCCs]) for (i) episodes with inclusion criteria, (ii) episodes without inclusion criteria, (iii) beneficiaries with inclusion criteria, and (iv) beneficiaries without inclusion criteria.

This analysis shows that the Renal or Ureteral Stone Surgical Treatment measure's inclusion criteria have only a minimal effect on the percentage of beneficiaries of any particular demographic. The difference between beneficiaries being included or not included in the measure is less than 2.3 percentage points across each of the characteristics in the analysis at TIN level testing, and less than 2.6 percentage points at TIN-NPI level testing. To illustrate, the percentage of beneficiaries aged 65 to 69 without applying the inclusion criteria is 30.0 percent, compared to 30.3 percent at both TIN and TIN-NPI level testing. The difference in the percentage of beneficiaries for race with and without the inclusion criteria is within one percentage point for all categories, with the exception of the White category, for which the difference in beneficiaries is 1.2 and 1.9 percentage points for TIN- and TIN-NPI level testing,

⁹ Only stays at IP facilities that are paid under 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. For details on the identification of these hospitals, please refer to the CCN definitions for Short-term (General and Specialty) Hospitals facility types in Chapter 2, Section 2779A1 of the [CMS State Operation Manual](#).

respectively. The breakdown of male and female beneficiaries remains almost the same when comparing the use of inclusion criteria at TIN-NPI level testing, with 42.8 percent female without inclusion criteria and 43.5 percent with inclusion criteria at both TIN and TIN-NPI levels. These results indicate that there is minimal shift in patient characteristics due to using the inclusion criteria listed above at both TIN and TIN-NPI level testing.

3.1.7 Sample Differences

n/a

3.1.8 Social Risk Factors Included in Analysis

The social risk factors analyzed were variables from the ACS, EDB, and CME. All ACS variables are at the Census Block Group level. Social risk variables analyzed include the following:

- Income (ACS)
 - Low Income: median income < 33rd percentile nationally
 - Medium Income: median income in the interval spanning the 33rd percentile to the 66th percentile nationally
 - High Income: median income > 66th percentile
- Education (ACS)
 - Education < High School: when % with < high school education is the highest for a given Census Block Group
 - Education = High School: when % with only high school is the highest
 - Education > High School: when % with > high school is the highest
- Employment (ACS)
 - Unemployment Rate > 10%
 - Unemployment Rate <= 10%
- Race (EDB)
 - Asian, Black, Hispanic, North American Native, White, and Other
- Sex (EDB)
 - Female, male
- Dual status (CME)
 - Full dual, partial dual, non-dual

3.2 Reliability Testing

3.2.1 Level of Reliability Testing

The following levels of reliability were tested: critical data elements used in the measure and performance measure score (e.g., signal-to-noise analysis).

3.2.2 Method of Reliability Testing

Data Element Reliability

The Renal or Ureteral Stone Surgical Treatment measure is constructed using CMS claims data, as described in Section 3.1.2. CMS has implemented several auditing programs to assess overall claims code accuracy, ensure appropriate billing, and recoup any overpayments. CMS routinely conducts data analysis to identify potential problem areas and detect fraud, and audits important data fields used in this measure, including diagnosis and procedure codes and other elements that are consequential to payment. Specifically, CMS works with Zone Program Integrity Contractors, and formerly Program Safeguard Contractors, to ensure program integrity; the agency also uses Recovery Audit Contractors to identify and correct for underpayments and overpayments.

CMS also uses the Comprehensive Error Rate Testing (CERT) Program to ensure that Medicare payments are correct in accordance with coverage, coding, and billing rules. Between 2005 and 2017, CERT estimates that proper payment, which includes payments that met Medicare coverage, coding, and billing rules, ranged from 87.3 to 96.4 percent of total payments each year.¹⁰ The fiscal year 2018 Medicare fee-for-service program proper payment rate was 91.9 percent.¹¹ CMS continues to perform successful corrective actions and give providers additional education to ensure accurate billing.

To ensure claims completeness and inclusion of any corrections, the measure was developed and tested using data with a three month claims run-out from the end of the measurement period.

Measure Reliability

Measure reliability is the degree to which repeated measurements of the same entity agree with each other. For measures of clinician performance, the measured entity is the TIN or TIN-NPI, and reliability is the extent to which repeated measurements of the TIN or TIN-NPI give similar results. To estimate measure reliability, we used a signal-to-noise analysis.

This approach seeks to determine the extent to which variation in the measure is due to true, underlying clinician performance rather than random variation (i.e., statistical noise) within clinicians due to the sample of cases observed. To achieve this, we calculate reliability scores as:

$$R_j = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_{w_j}^2}$$

Where:

$\sigma_{w_j}^2$ is the within-group variance of the mean measure score of clinician j

σ_b^2 is the between-group variance of clinicians within the episode group

That is, reliability is calculated as the ratio of between-group variance to the sum of between-group variance and within-group variance. Reliability closer to a value of one indicates that the between-group variance is relatively large compared to the within-group variance, which suggests that the measure is effectively capturing the systematic differences between the clinician and their peer cohort.

3.2.3 Statistical Results from Reliability Testing

Measure Reliability

As displayed in the table below, 100 percent of TINs and TIN-NPIs at 10, 20, and 30-episode volume thresholds have mean reliability greater than or equal to 0.4. At a testing volume threshold of at least 10 episodes, the mean reliability for TINs is 0.77 and for TIN-NPIs is 0.65. The mean reliability continues to increase at the 20 and 30-episode volume thresholds.

¹⁰ Comprehensive Error Rate Testing (CERT) Program. "Appendices Medicare Fee-for-Service 2018 Improper Payments Report". Table A6. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/Medicare-FFS-Compliance-Programs/CERT/Downloads/2018MedicareFFSSupplementalImproperPaymentData.pdf>.

¹¹ Ibid.

Table 2: Reliability Results at Various Volume Thresholds

Volume Threshold (# episodes)	TIN		TIN-NPI	
	Mean Reliability	% ≥ 0.4	Mean Reliability	% ≥ 0.4
10	0.77	100.0%	0.65	100.0%
20	0.84	100.0%	0.75	100.0%
30	0.87	100.0%	0.80	100.0%

3.2.4 Interpretation

Measure Reliability

Overall reliability of the Renal or Ureteral Stone Surgical Treatment measure exceeds the CMS reliability threshold at a volume threshold of 10 episodes or more for both TINs and TIN-NPIs due to the large number of episodes attributed to clinicians. CMS generally considers 0.4 as the threshold indicating ‘moderate’ reliability, which is supported by previous work into reliability.¹²

While higher volume thresholds yield even higher reliability results, it is at the cost of further reducing the number of clinicians and clinician groups able to receive a measure score.

3.3 Validity Testing

3.3.1 Level of Validity Testing

We conducted performance measure score validity testing, which included systematic assessment of face validity and empirical validity testing.

3.3.2 Method of Validity Testing

Face Validity

The Renal or Ureteral Stone Surgical Treatment measure was developed through a structured, iterative process for gathering detailed input from recognized clinician experts on the measure. These expert panels were convened to methodically assess the extent to which the measure: (i) captured what it was intended to capture, and (ii) differentiated between provider performance. Experts in this clinical area evaluated specifications in an iterative process to ensure that each aspect of the measure (e.g., assigned services) was intentionally capturing only the costs of care within the reasonable influence of the attributed clinician for a defined patient population (i.e., the ability of the measure score to differentiate good from poor performance).

In developing and refining this measure, Acumen incorporated input from (i) the Urologic Disease Management Clinical Subcommittee, (ii) the Renal or Ureteral Stone Surgical Treatment workgroup, (iii) a Technical Expert Panel (TEP), (iv) a Person and Family Committee (PFC), and (v) stakeholder feedback from national field testing.

The Clinical Subcommittee comprised 24 members with clinical experience in urologic disease management, affiliated with 22 specialty societies. The Clinical Subcommittee provided input at an in-person meeting in April 2018 on which measure to develop, on the measure scope, and on the composition of a smaller, targeted workgroup to provide detailed input on each aspect of measure specifications. The Renal or Ureteral Stone Surgical Treatment workgroup was composed of 12 members affiliated with 9 specialty societies, including the Renal Physicians Association, the American Urological Association, the American Association of Clinical

¹² Mathematica, Inc., “Memorandum: Reporting Period and Reliability of AHRQ, CMS 30-Day and HAC Quality Measures – Revised,” http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/Downloads/HVBP_Measure_Reliability-.pdf.

Urologists, and the Society of Interventional Radiology. The workgroup considered empirical analyses and their clinical expertise to provide input during an in-person meeting and several webinars between June and December 2018. Input was gathered in a structured manner including the use of a polling process requiring greater than 60 percent consensus.

The TEP provided high-level guidance and input on the overall direction of measure development and the framework for episode-based cost measures, while the PFC provided a patient and caregiver perspective. PFC input included concepts of healthcare quality and value, guiding principles and measure-specific topics to inform the workgroups such as pre- and post-trigger windows for selected episodes, and inclusion of services and costs for attributed clinicians. In addition, the national field testing feedback period in October and November 2018 offered all stakeholders an opportunity to review and provide input on draft measure specifications and measure feedback reports for attributed clinicians and clinician groups. During this period, 78,221 field test reports for TINs and TIN-NPIs were available for download and review for 11 episode-based cost measures developed throughout 2018.

One of the key roles of the measure-specific workgroup was to develop service assignment rules for the cost measure. These service assignment rules are intended to ensure clinicians are evaluated on services and costs that are clinically related to the attributed clinician's role in surgical treatment for renal or ureteral stone, thus preventing inclusion of unrelated cost variation in this measure. Assigned services occurring in the outpatient setting was defined separately for the pre- and post-trigger windows, and includes renal or ureteral stone surgery, evaluation, imaging, testing, treatment, complications, and follow-up.

Empirical Validity Testing

We undertook two approaches to estimate the measure's validity. In the first approach, we evaluated the empirical validity of the Renal or Ureteral Stone Surgical Treatment measure by examining differences in risk-adjusted cost for known indicators of resource or service utilization based on a literature review, specifically complications related to surgical treatment for renal or ureteral stone. For this analysis, we compared the observed over expected (O/E) cost ratio for Renal or Ureteral Stone Surgical Treatment episodes with and without complications related to procedures occurring in the post-trigger period. This analysis sought to confirm the expectation that the Renal or Ureteral Stone Surgical Treatment measure captures variation in service utilization.

In the second approach, we evaluated how different types of cost impact risk-adjusted measure scores. Certain services or costs included in the Renal or Ureteral Stone Surgical Treatment measure were classified into clinically coherent groups of services, called "clinical themes." The Renal or Ureteral Stone Surgical Treatment measure clinical themes are:

- **Preoperative Management:** Imaging, testing, or transurethral excision, drainage or removal for obstructive and reflux uropathy or calculus of kidney and ureter.
- **Postoperative Management:** Medications, testing and procedures related to kidney or ureter disorders, complications from implants, devices or grafts.
- **Rehabilitation / Durable Medical Equipment (DME) / Supplies:** Including, walkers, catheters, drainage bags, ostomy pouches, dressing, wound care, and wheelchairs.
- **Preoperative Stent or Catheter Placement:** Catheterization, nephrostomy, extracorporeal lithotripsy, or other therapeutic procedures of the urinary tract.
- **Postoperative Stent Placement / Removal / Exchange:** Catheterization and removal or replacement of stent or drainage tube.
- **Follow-Up Visit Related to Postoperative Pain:** Diagnostic and therapeutic procedures related to abdominal or pelvic pain, nausea and vomiting, or general pain and related medications and supplies.

- **Postoperative Infection, Other:** Care for sepsis, urinary system disorders, fever, or bacterial infection, including diagnostic and therapeutic procedures, medications, biopsy, and related supplies.
- **Postoperative UTI or Procedure-Related Infection:** Care for cystitis, pyelonephritis, or other disorders of the urinary system (such as urinary tract infections [UTIs]), and complications of devices, implants or grafts, including diagnostic and therapeutic procedures, medications, transportation, biopsy, transfusions, wound care catheterization, and related supplies.
- **Repeat Procedure for Stones:** Care related to kidney stones, including transurethral excision, drainage or removal for obstruction, testing, dilation, or extirpation of matter.
- **Other ER Visit or Hospitalization:** Kidney and ureter procedures for non-neoplasm, ultrasonography, transfusions, or fluoroscopies.

As with the first analysis for validity, the aim of this analysis was to determine whether the measure is capturing variation in provider cost in the manner intended and expected. To measure this, we took the Pearson correlation between the cost of each clinical theme and the overall risk-adjusted cost for an episode.

We expected that the Postoperative Infection, Other theme would have the highest correlation with risk-adjusted episode cost, as non-UTI complications such as sepsis are likely associated with high cost even after accounting for beneficiary characteristics.¹³ We would expect similar trends for the Postoperative UTI or Procedure-Related Infection theme as it also contains services relating to complications. By contrast, we expected that Follow-Up Visit Related to Postoperative Pain to have smaller correlations to risk-adjusted cost, as services under this theme are more less commonly associated with patient complexity or procedure-related complications, both of which are common drivers of cost.

3.3.3 Statistical Results from Validity Testing

For the first analysis of validity, the mean O/E cost ratio for all episodes is 1.00. The mean O/E cost ratio for episodes with services relating to complications during the post-trigger period is 1.14, compared with 0.92 for episodes without services relating to complications during the post-trigger period. Table 3 contains additional details on the O/E cost ratios for the various episode types.

Table 3: Distribution of Observed to Expected Cost Ratios

Episode Type	Observed to Expected Cost Ratio										
	Mean	Std. Dev.	Percentile								
			1st	5th	10th	25th	50th	75th	90th	95th	99th
All Final Episodes	1.00	0.36	0.28	0.64	0.71	0.80	0.89	1.12	1.50	1.71	2.29
Episodes with Services Related to Stone Surgery Complications	1.14	0.41	0.44	0.70	0.77	0.85	0.99	1.35	1.69	1.95	2.52
Episodes without Services Related to Stone Surgery Complications	0.92	0.30	0.22	0.62	0.69	0.77	0.85	0.98	1.31	1.53	2.05

¹³ Khan, N.A., Quan, H., et al., "Association of postoperative complications with hospital costs and length of stay in a tertiary care center" J Gen Intern Med (2006) 21: 177.

The clinical themes analysis demonstrates that there is a strong correlation between the Postoperative UTI or Procedure-Related Infection (correlation: 0.67) and Preoperative Stent or Catheter Placement (correlation: 0.47) themes and risk-adjusted cost. By contrast, the Postoperative Stent Placement / Removal / Exchange (correlation: 0.04) and Preoperative Management (correlation: 0.22) themes had lower correlation with risk-adjusted cost.

3.3.4 Interpretation

As expected, the average O/E cost ratios for episodes with and without services related to complications are close to one, demonstrating that the Renal or Ureteral Stone Surgical Treatment measure is able to accurately capture costs across a very wide range of resource use.

The clinical themes analysis demonstrates that high risk-adjusted cost is strongly associated with themes related to complications, such as Postoperative UTI or Procedure-Related Infection, and also linked—though more weakly, as expected—to themes relating to follow-up care and preoperative management, such as Preoperative and Postoperative Management. This indicates that the measure may penalize clinicians who have higher rates of complications, while not disincentivizing the provision of appropriate pre- and post-operative care, such as pain management. Importantly, we see that correlation with risk-adjusted cost is similarly strong for themes with substantial cost differences, such as Postoperative UTI or Procedure-Related Infection (average cost: \$1,689.85; correlation: 0.67) and Postoperative Infection (average cost: \$3,142.82; correlation: 0.65). This indicates that the correlation does not come from a mechanical increase in episode costs from high-cost themes.

3.4 Exclusions Analysis

3.4.1 Method of Testing Exclusions

Exclusions are used in the Renal or Ureteral Stone Surgical Treatment measure to ensure a homogenous patient population within the scope of the measure focus on surgical procedures for renal or ureteral stone treatment and that episodes provide meaningful information to attributed clinicians or as part of data processing, to ensure that sufficient data are available to accurately determine episode spending and calculate risk adjustment for each episode. For the exclusions analysis, we focused on exclusions added to ensure a homogenous patient population. These exclusions, along with their rationales, are listed below:

- *Episodes where beneficiary death date occurred before the episode end.*
 - These episodes are excluded for all measures due to the potential to inaccurately reflect a clinician's performance. Episodes where the beneficiary died may be unusually high-cost, due to perimortem treatment costs, or unusually low-cost, due to the truncated episode window. Neither of these cases accurately reflects the efficiency of the clinician performing the treatment.
- *Episodes where the beneficiary has or had a stone of the lower urinary tract or bladder.*
 - These patients require clinically distinct treatment and represent a separate, smaller, and less uniform population, and therefore should not be included in this measure.
- *Episodes with any additional non-trigger, non-endoscopic OP urologic procedures.*
 - Patients undergoing more than one urologic procedure at a given time are nonstandard and therefore could not be compared fairly to the broader patient population.
- *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.

Given the rationales for these exclusions, we would expect these excluded episodes to have a different risk profile than the included episodes, such as a higher mean cost or a different distribution of costs (e.g., a long tail of high-cost episodes). For the exclusions, we examined the number of episodes and beneficiaries affected, as well as the distributions of observed cost and O/E cost ratio (calculated by applying existing risk factor coefficients to the excluded episodes) for excluded episodes. We then compared the cost characteristics of the excluded episodes to those of final episodes included in measure calculation to assess the distinctness between the two patient cohorts. A full list of the exclusions and details used for the Renal or Ureteral Stone Surgical Treatment measure is provided in the Measure Codes List.¹⁴

3.4.2 Statistical Results from Testing Exclusions

Table 4 below presents observed cost statistics and O/E cost ratios for the Renal or Ureteral Stone Surgical Treatment measure exclusions. Cost statistics are also provided for the set of final episodes included in the Renal or Ureteral Stone Surgical Treatment measure for comparison, with a testing volume threshold of 10 episodes at the TIN and TIN-NPI levels.

Table 4: Cost Statistics for Measure Exclusions

Exclusion	Episodes		Observed Cost			O/E		
	#	%	Mean	Percentile		Mean	Percentile	
				10 th	90 th		10 th	90 th
All Episodes Meeting Triggering Logic	129,316	100.00%	\$6,430	\$2,922	\$11,123	1.00	0.63	1.53
Beneficiary Death in Episode	1,713	1.32%	\$7,908	\$1,275	\$18,105	1.05	0.23	1.97
Non-trigger/Non-endoscopic Urologic Procedures	33	0.03%	\$3,076	\$575	\$6,451	0.47	0.10	1.20
Stone of Lower Urinary Tract/Bladder	11,550	8.93%	\$6,776	\$2,576	\$12,352	1.00	0.43	1.61
Outlier	2,100	1.62%	\$13,600	\$811	\$26,259	1.66	0.13	3.46
<i>Final Episodes (TIN)</i>	99,613	77.03%	\$6,171	\$3,477	\$10,036	0.99	0.70	1.48
<i>Final Episodes (TIN-NPI)</i>	85,207	65.89%	\$6,178	\$3,456	\$10,076	0.99	0.70	1.48

3.4.3 Interpretation

These episodes were excluded due to clinical considerations to ensure a comparable patient cohort that will yield meaningful information to attributed clinicians. Further discussion of the results for each exclusion is provided below.

Episodes ending in death: While there is considerable difference between mean observed episode cost for episodes ending in death and the final set of episodes (\$7,908 compared to \$6,171 at the TIN level), the distribution of costs for episodes ending in death is substantially wider. At the 10th percentile, the episode cost is \$1,275, less than half of the cost of the final set of episodes and at the 90th percentile, is nearly double (\$18,105 compared to \$10,036 for the final set of episodes at the TIN level). This reflects the two scenarios discussed above: that episodes ending in death can appear low cost due to the truncated episodes or much more expensive due to costly services during the time leading up to a beneficiary's death.

¹⁴ CMS, "Renal or Ureteral Stone Surgical Treatment Measure Codes List," *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2019-revised-ebcm-measure-specs.zip>.

Episodes where the beneficiary has or had a stone of the lower urinary tract or bladder: The mean observed cost of these episodes was on average slightly more expensive than the final-episode mean (\$6,776 compared to \$6,171 at the TIN level). Urologic stones in the lower urinary tract or bladder are treated differently, and would not be clinically comparable to the treatment of renal stones.

Episodes with any additional non-trigger, non-endoscopic OP urologic procedures: The mean observed cost of these episodes is less than half the mean cost of the final episodes (\$3,076 compared to \$6,171 at the TIN level). This very small patient population (33 episodes) are atypical relative to the broader cohort undergoing the procedure.

Outlier cases: The O/E cost ratio ranges from 0.13 at the 10th percentile to 3.46 at the 90th percentile, indicating that the risk adjustment model is currently unable to account for the patient characteristics associated with these high- and low-cost outlier episodes. Excluding outliers based on risk-adjusted cost eliminates the episodes that deviate most from expected spending levels based on patient characteristics.

3.5 Risk Adjustment or Stratification

3.5.1 Method of Controlling for Differences

Differences in case mix are controlled for using a statistical risk model with 107 risk factors and stratification by 3 risk categories.

The risk adjustment model for the Renal or Ureteral Stone Surgical Treatment measure broadly follows the CMS-HCC risk adjustment methodology, which is derived from Medicare Parts A and B claims and is used in the Medicare Advantage (MA) program. Although the MA risk adjustment model includes 24 age/sex variables, this risk adjustment model does not adjust for sex and so only includes 12 age categorical variables. Severity of illness is measured using HCCs, indicators of enrollment and long-term care status, and disease interactions. The risk adjustment model also includes variables for factors identified by the expert clinician workgroup as affecting resource use.

The model includes 79 HCC indicators derived from the beneficiary's Parts A and B claims during the period 120 days prior to the episode trigger and are specified in the CMS-HCC Version 22 (V22) 2016 model. Episodes for beneficiaries without a full 120-day lookback period are excluded from the measure. This 120-day period is used to measure beneficiary health status and ensures that each beneficiary's claims record contains sufficient fee-for-service data both for measuring spending levels and for risk adjustment purposes.

In addition, the risk adjustment model includes status indicator variables for whether the beneficiary qualifies for Medicare through Disability or ESRD. The model also includes an indicator of whether the beneficiary recently required long-term care, defined as 90 days in a long-term care facility without being discharged to community for 14 days. Beneficiaries who need to reside in long-term care facilities typically require more intensive care than beneficiaries who live in the community. These enrollment and long-term care status variables are non-diagnostic indicators of severity of illness.

The model also accounts for disease interactions between HCCs and/or enrollment status variables included in the MA model. These interactions are included because certain combinations of comorbidities increase costs more than is predicted by the HCC indicators alone.

Furthermore, the risk adjustment model includes measure-specific factors intended to further isolate costs that attributed clinicians can reasonably influence, informed by expert clinician

input and empirical analyses. The following variables were added to avoid potential unintended consequences:

- whether the procedure was performed in an ASC or HOPD setting, since costs may differ and choosing place of service may not be under the physician's control;
- whether the procedure was contralateral;
- whether the beneficiary received a planned staged/repeat unilateral procedure, because costs will be inherently higher for a staged unilateral procedure than a single unilateral, and because unplanned procedures differ from planned procedures in that they likely represent management of complication(s);
- whether the trigger service is the second planned staged/repeat unilateral procedure since costs will be different for a staged unilateral procedure than single unilateral, and planned procedures differ from unplanned procedures in that the latter is likely management of complication(s);
- whether the trigger service is the second unplanned staged/repeat unilateral procedure since costs will be inherently higher for a staged unilateral procedure than single unilateral, and unplanned procedures differ from planned procedures in that they likely represent management of complication(s);
- whether the beneficiary has a history of opiate dependence or chronic pain, possibly requiring more complex treatment, e.g., with regards to pain management;
- whether the trigger service is for treatment of a ureteral stone with obstruction, because such patients may require additional services; and
- whether the trigger service is for treatment of a urinary stone with UTI or renal infection, since patients with associated infection inherently require additional services.

As with the CMS-HCC model, the risk adjustment approach for this measure uses an ordinary least squares linear regression model. The predicted, or expected, cost is winsorized at 0.5th percentile to make sure episodes with unusually small predicted cost, which would lead to abnormally large O/E ratios, do not dominate certain clinicians' final score. The winsorized expected costs are renormalized to ensure the average expected episode cost is the same before and after winsorizing. Then, as noted in the exclusions analysis above, extremely low- or high-cost outlier episodes with residuals below the 1st percentile or above the 99th percentile are excluded to reduce the effect of episodes that deviate the most from their expected values in absolute terms. The expected cost after excluding these outliers is again renormalized to ensure that average expected costs are the same after outlier removal.

Finally, the risk adjustment model outlined above is performed separately for each of the three Renal or Ureteral Stone Surgical Treatment measure sub-groups, which are based on the technique used to treat the stone.

- Extracorporeal Shock Wave Lithotripsy (ESWL)
- Percutaneous Nephrostolithotomy
- Ureteroscopy

Full details of the risk adjustment model are in the Measure Codes List File.¹⁵ The National Summary Data Report (NSDR) Addendum includes regression coefficients and standard errors

¹⁵ CMS, "Renal or Ureteral Stone Surgical Treatment Measure Codes List," *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2019-revised-ebcm-measure-specs.zip>.

for each of the covariates used in the risk adjustment model.¹⁶

3.5.2 Conceptual, Clinical, and Statistical Methods

We selected the CMS-HCC model based on previous studies evaluating its appropriateness for use in risk adjusting Medicare claims data. This model was developed specifically for use in the Medicare population, meaning that it accounts for conditions found in the Medicare population and is calibrated on Medicare fee-for-service beneficiaries. In addition, the CMS-HCC model is routinely updated for changes in coding practices (e.g., the transition from ICD-9 to ICD-10 codes) and is exhaustive on these code sets. Because the CMS-HCC model has already been extensively tested, we focus our testing on how the CMS-HCC model was adapted to the Renal or Ureteral Stone Surgical Treatment measure methodology.

The workgroup provided input on measure-specific risk adjusters after reviewing empirical analyses on subpopulations of interest to assess whether and if so, how, particular factors should be accounted for in the model. These could include patient characteristics, factors outside of the attributed clinicians' influence, or any other factors that would help prevent unintended consequences. These additional risk adjusters are listed in the section above.

As previously noted, the risk adjustment model is run on episodes stratified into sub-groups which may qualify as "ordering" of risk factors. Sub-groups were also determined based the workgroup's input, with the goal of ensuring clinical comparability among episodes so that the cost measure fairly compares clinicians with similar patient case-mix. The sub-groups, which are based on the procedural approach used, are listed in the above section. Each of these three procedures are unique approaches to management of urinary system stone, which may require a different set of preoperative services and have different risk for postoperative services. Percutaneous nephrostolithotomies are the more invasive procedures since they involve skin incisions and the endoscope can approach the ureter or kidney from the peritoneum or retroperitoneum.

3.5.3 Conceptual Model of Impact of Social Risks

Our conceptual model of the impact of social risk factors is informed by both published, peer-reviewed literature and data analysis.

3.5.4 Statistical Results

The literature has extensively tested the use of the HCC model as applied to Medicare claims data. Although the variables in the HCC model were chosen to predict annual cost, CMS has also used this risk adjustment model in a number of other settings (e.g., ACOs, previous physician QRUR programs, and other measures such as National Quality Forum (NQF) #2158: MSPB-Hospital cost measure). Recalling that the risk model relies on the existing CMS-HCC model, testing results for factors included in the CMS-HCC V22 2016 model can be found in the Pope et al (2011) report.¹⁷ For measure-specific factors not included in the CMS-HCC model, we sought expert clinician input through the workgroup, which provided recommendations on additional risk adjusters and sub-groups.

¹⁶ CMS, "National Summary Data Report Addendum: 11 Episode-Based Cost Measures and Revised MSPB Clinician Measure," *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html>.

¹⁷ Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

The results of the statistical analysis used to characterize our risk adjustment model can be found in the NSDR Addendum, which includes regression coefficients and standard errors for each of the covariates used in the risk adjustment model.

3.5.5 Analyses and Interpretation in Selection of Social Risk Factors

Acumen analyzed gender, dual status, income, education, and unemployment as social risk factors (more information on these variables can be found in Section 3.1.8). Beneficiary gender and dual status were obtained from the EDB and CME. Information on income, education, and unemployment was obtained from ACS data and linked to episodes by census block group where possible to provide a more granular level of analysis than ZIP code.

The percentage of female beneficiaries range from 42.2 percent to 50.8 percent across the three sub-groups in this measure. The majority of the beneficiaries (78.57% – 86.63%) have non-dual status. Income level is categorized into high, medium, and low from the continuous average income variable in ACS; therefore, each category has 33.3 percent of observations. While 1.7 to 1.9 percent of beneficiaries are classified below a high school education level, over 98 percent of episodes are classified at a high school level or greater. Finally, 20.7 to 22.0 percent of beneficiaries have high unemployment designation (>10%).

Acumen examined the impact of including social risk factors into our risk adjustment model by running goodness of fit tests when different risk factors are added and compared to the base risk adjustment model, where the base risk adjustment model refers to the full standard set of risk adjustment variables from the CMS-HCC V22 2016 model, disability status, ESRD status, interaction variables, recent long-term care use, and measure-specific clinical risk adjusters. Acumen ran a step-wise regression to include gender, dual status, gender + dual status, and gender + dual + income + education + unemployment + race, on top of the adapted CMS-HCC model. The step-wise regressions help evaluate individual as well as joint significance of the social risk factors. We examined the impact of including social risk factors into our risk adjustment model with T-test of individual significance and F-test of joint significance.

First, we analyzed the model coefficients and p-values for each of the base and social risk factor models to understand whether any of the social risk factor covariates are predictive of episode cost. The T-test and F-test revealed many significant p-values, indicating that social risk factors are likely predictive factors for determining resource use among beneficiaries for the relevant characteristic. However, the analysis also shows that the directions of the effects of social risk factors are not consistent. For example, morbidly obese beneficiaries' episodes may display higher spending for the Percutaneous Nephrostolithotomy sub-group but lower spending for the other two sub-groups.

Secondly, we analyzed the impact of adding social risk variables on overall model performance by looking at the differences in the O/E cost ratios with and without social factors in the risk adjustment model. When including social risk factors in our risk adjustment regression, the minor differences in the O/E ratios, even for providers at high or low extremes of risk, indicates that social risk factor effects on the model performance are likely captured through existing risk adjustment variables. When including the social risk factors in risk adjustment, the O/E cost ratio for 94.9 percent of TINs and 92.9 percent of TIN-NPIs changed by ± 0.01 or less.

Finally, we analyzed the correlation between measure scores calculated with and without the social risk factors. The measure scores calculated with and without these social factors were highly correlated at both the TIN and TIN-NPI levels, with a Spearman correlation coefficient of 0.999 at both levels. These results indicate that the inclusion of social risk factors in the current risk adjustment model would have a limited effect on measure scores.

Due to the inconsistent direction and limited impact of social risk factor effects under the current risk adjustment model, we believe the Renal or Ureteral Stone Surgical Treatment measure risk adjustment model sufficiently accounts for the effects of social risk factor on clinician measure scores.

3.5.6 Method for Statistical Model or Stratification Development

To analyze the validity of current risk adjustment model, we examined three analyses: (1) R-squared and adjusted R-squared for the regression models, (2) predictive ratios and o/e cost ratios to examine the fit of the models at different levels of patient complexity, and (3) coefficient estimates, standard errors, and p-values for each sub-group.

- 1) *R-squared and adjusted R-squared* were calculated for the measure overall as well as for each sub-group. The results should be evaluated in the context of the service assignment rules, which indicate which costs are counted in the measures and which costs are not counted. This is an important distinction from all-cost measures, as a low R-squared does not necessarily indicate that a measure reflects variation unrelated to clinical care, while a high R-squared does not necessarily indicate the opposite; instead, the risk adjustment models must be evaluated in concert with the service assignment rules. These results are provided in Section 3.5.7.
- 2) *Predictive ratios and O/E cost ratios* were calculated for each “risk decile” for the episode group. A “risk decile” is based on the risk scores, which indicate how costly episodes are expected to be, as predicted through risk adjustment. After arranging episodes into deciles based on their risk score, we calculated the predictive ratios and average O/E cost ratios for each decile. The predictive ratio aims to examine the fit of the model at different levels of patient complexity to examine the model’s ability to predict both very low and high cost episodes, and is calculated using the formula of average (expected cost)/average (observed cost) for all episodes in each decile. Similarly, the O/E cost ratio demonstrates the model’s prediction accuracy, and is calculated using the formula of average (observed cost/expected cost) for all episodes in each decile. These are discussed in Sections 3.5.8 and 3.5.9.
- 3) *Coefficient estimates, standard errors, and p-values* were run for each sub-group to consider the extent to which the coefficients for the risk factor covariates are predictive of episode cost. Results for individual risk adjustment variables should be viewed in the context of the entire model and set of sub-groups, rather than being analyzed individually. For instance, coefficients indicate the incremental effect of a model variable, holding all other variables fixed. As another example, interactions between model variables must be interpreted in concert with the effects of those variables in isolation.

The results of these analyses are presented in the NSDR Addendum to aid in the overall assessment of the predictive ability of the risk adjustment models.¹⁸

3.5.7 Statistical Risk Model Discrimination Statistics

The overall R-squared for the Renal or Ureteral Stone Surgical Treatment cost measure, calculated by dividing explained sum of squares by total sum of squares is 0.46. The adjusted R-squared is 0.46.

¹⁸ CMS, “National Summary Data Report Addendum: 11 Episode-Based Cost Measures and Revised MSPB Clinician Measure,” *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html>.

The NSDR Addendum also includes regression coefficients and standard errors for each of the covariates used in the risk adjustment model. More information on discrimination testing for the CMS-HCC model can be found at Pope et al. 2011.¹⁹

3.5.8 Statistical Risk Model Calibration Statistics

We interpret calibration as how accurately the risk model's predictions match the actual episode cost. We calculate the average O/E cost ratio for each risk decile to demonstrate the model's prediction accuracy. The average O/E cost ratio is generally close to one across risk deciles, indicating that the model is accurately predicting actual episode cost. Full results can be seen the NSDR Addendum.

3.5.9 Statistical Risk Model Calibration – Risk Decile

Analysis of predictive ratios by risk decile for the measure shows that the model has consistent predictive ratios across risk score deciles, with each decile having a predictive ratio between 0.98 and 1.01.

3.5.10 Results of Risk Stratification Analysis

Results indicate that the three measure sub-groups have varying measure scores (see below table). Specifically, Percutaneous Nephrostolithotomy cases are more expensive on average than ESWL or Ureteroscopy cases. At the TIN level, the mean score for Percutaneous Nephrostolithotomy episodes is \$11,023, compared to ESWL episodes at \$5,654, and Ureteroscopy cases at \$6,350. Results are similar at the TIN-NPI level. Thus, cases are considered separately based upon the surgical approach so as not to create financial incentives that might influence procedure choice.

Table 5: Distribution of Score by Sub-Group

Level	Sub-Group	Provider Count	Mean Score	Score Percentile						
				1st	10th	25th	50th	75th	90th	99th
TIN	All TINs	1,661	\$6,192	\$4,836	\$5,422	\$5,730	\$6,082	\$6,545	\$7,123	\$8,519
TIN	Percutaneous Nephrostolithotomy	1,535	\$5,654	\$3,959	\$4,760	\$5,102	\$5,489	\$6,028	\$6,731	\$8,810
TIN	Extracorporeal Shock Wave Lithotripsy	709	\$11,023	\$2,415	\$8,453	\$9,733	\$10,769	\$12,226	\$14,249	\$19,325
TIN	Ureteroscopy	1,630	\$6,350	\$4,485	\$5,332	\$5,743	\$6,215	\$6,786	\$7,535	\$9,677
TIN-NPI	All TIN-NPIs	4,158	\$6,170	\$4,697	\$5,272	\$5,635	\$6,074	\$6,595	\$7,166	\$8,727
TIN-NPI	Percutaneous Nephrostolithotomy	3,631	\$5,605	\$3,748	\$4,580	\$4,931	\$5,399	\$6,050	\$6,860	\$9,424
TIN-NPI	Extracorporeal Shock Wave Lithotripsy	1,187	\$11,062	\$2,028	\$8,279	\$9,693	\$10,755	\$12,362	\$14,555	\$21,922
TIN-NPI	Ureteroscopy	4,031	\$6,309	\$4,231	\$5,162	\$5,592	\$6,158	\$6,852	\$7,614	\$9,722

¹⁹ Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

3.5.11 Interpretation

The R-squared values for the model, which measure the percentage of variation in results predicted by the model, are higher than the values presented in similar analyses of risk adjustment models.²⁰ As noted in Section 3.5.6, these results should be interpreted alongside service assignment rules, which remove clinically unrelated services, so the resulting variation is reflective of variation related to factors within a clinician's reasonable influence.

As demonstrated in Sections 3.5.8 and 3.5.9, the average O/E cost ratios and the predictive ratios for all risk deciles are close to one. Predictive ratios close to one indicate that expected spending is accurately predicting observed spending. Overall, the results show that the model is accurately predicting observed spending, regardless of overall risk level.

3.6 Identification of Meaningful Differences in Performance

3.6.1 Method

Our method of determining clinically meaningful differences in episode-based cost measure scores consists of stratifying the clinician measure scores by meaningful characteristics and investigating the clinician score distribution by percentile. Stratification is performed for each of the following characteristics: urban/rural, census division, census region, risk score, and the number of episodes attributed to the clinician. We analyze the distribution of measure scores for clinicians defined by these characteristics, as well as for the overall episode group and for each sub-group.

The purpose of this analysis is to ensure that there is a sufficiently large difference in measure scores among clinicians to determine a meaningful difference in performance. In addition, this analysis looks to confirm that the measure behaves as expected with respect to meaningful clinician characteristics.

3.6.2 Statistical Results

Key findings show that, generally, there is a large performance difference among clinicians in the Renal or Ureteral Stone Surgical Treatment measure:

- (i) the 99th percentile of the measure score is approximately 1.8 times the 1st percentile at both the TIN level and TIN-NPI levels;
- (ii) the Renal or Ureteral Stone Surgical Treatment measure score at the 90th percentile is approximately 31 percent greater than the score at the 10th percentile at the TIN Level and approximately 35 percent greater at the TIN-NPI level; and

The results also show that there is not systemic regional difference in clinician score. For instance, the mean scores for clinicians across nine census divisions (excluding 'Unknown') are within a less than 8% range (i.e., \$5,994 – \$6,454 at the TIN level and \$5,940 – \$6,331 at the TIN-NPI level). Similarly, clinicians in urban areas seem to perform comparably to those in rural areas (less than \$200 difference at TIN and TIN-NPI levels).

In terms of other clinician characteristics, analysis of clinicians by number of episodes indicates that clinicians with more episodes perform similarly to those who perform fewer surgical treatments for renal or ureteral stones (under \$250 difference at both TIN and TIN-NPI levels). We also analyzed clinicians by risk score decile, as variation by risk score decile could indicate that the risk adjustment model is over- or under-correcting for clinicians with systematically

²⁰ Pope, Gregory C., John Kautter, Melvin J. Ingber, Sara Freeman, Rishi Sekar, and Cordon Newhart. "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

riskier patients. The mean measure scores also show little variation by risk score decile, with a range of \$5,972 to \$6,290 for TINs and a range of \$5,973 to \$6,325 for TIN-NPIs, indicating that the risk adjustment model is overall functioning as intended. Full results can be seen in the NSDR.²¹

3.6.3 Interpretation

There is clinically and practically significant variation in Renal or Ureteral Stone Surgical Treatment measure scores, indicating the measure's ability to capture differences in performance. Our findings regarding variation in measure scores are consistent with expert clinician input. The Renal or Ureteral Stone Surgical Treatment measure-specific workgroup suggested development of sub-groups based on surgical method used, noting the differences in cost between different procedures. Overall, as expected, results show that clinicians are not being systematically penalized or rewarded due to risk score decile given the current Renal or Ureteral Stone Surgical Treatment measure design (i.e., the differences in performance are not due to the risk profile of patients).

3.7 Missing Data Analysis and Minimizing Bias

3.7.1 Method

Since CMS uses Medicare claims data to calculate the Renal or Ureteral Stone Surgical Treatment measure, Acumen expects a high degree of data completeness. To ensure that we have complete and accurate data for each beneficiary who opens an episode, Acumen excludes episodes where beneficiary date of birth information (an input to the risk adjustment model) cannot be found in the EDB, the beneficiary does not appear in the EDB, or the beneficiary death date occurs before the episode trigger date.

The Renal or Ureteral Stone Surgical Treatment measure also excludes episodes where the beneficiary is enrolled in Medicare Part C or has a primary payer other than Medicare in the 120-day lookback period and episode window. In such situations, Medicare Parts A and B claims data may not capture the complete clinical profile for the beneficiary needed to capture the clinical risk of the beneficiary in risk adjustment. Furthermore, Parts A and B claims data may not capture all Medicare resource use if some portion of the beneficiary's care is covered under Medicare Part C.

3.7.2 Missing Data Analysis

The table below presents the frequency of missing data across the four categories of missing data, which caused episodes to be excluded from the Renal or Ureteral Stone Surgical Treatment measure. Frequency is presented in terms of the number of episodes excluded due to missing data, as well as the number of TINs and TIN-NPIs who had at least one episode excluded due to missing data. The missing data categories are:

- Beneficiary date of birth is missing
- Beneficiary death date occurred before the trigger date
- Beneficiary has a primary payer other than Medicare during the episode window or in the 120-day lookback period
- Beneficiary was not enrolled in Medicare Parts A and B, or was enrolled in Part C, during the 120-day lookback period and episode window

²¹ CMS, "National Summary Data Report: 11 Episode-Based Cost Measures and Two Revised Cost Measures, Updated Following Field Testing (Oct-Nov 2018)," *MACRA Feedback Page*, <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/macra-feedback.html>.

Table 6: Missing Data Categories for the Renal or Ureteral Stone Surgical Treatment Measure

Exclusion	# Episodes	# TINs	# TIN-NPIs
Missing Beneficiary Birth Date	0	0	0
Primary Payer Other than Medicare	15,491	1,907	5,594
Beneficiary Death before Admission	*	*	*
No Continuous Enrollment in Medicare Parts A and B, and Any Enrollment in Part C	7,548	1,633	4,036

*denotes that there were fewer than 11 episodes

3.7.3 Interpretation

As the Renal or Ureteral Stone Surgical Treatment, measure is calculated with Medicare claims data, Acumen expects a high degree of data completeness, which is supported by the limited frequency of missing data as noted above. Acumen takes measures to ensure that missing or inaccurate information in claims data is not included in the cost measure.

4.0 Feasibility

4.1 Data Elements Generated as Byproduct of Care Processes

The data elements used in this measure are generated, collected, and/or used by healthcare personnel during the provision of care (e.g., blood pressure, laboratory values, diagnosis, depression score). The data collected during care provision are then translated into the appropriate coding system (e.g. ICD-10 diagnoses, MS-DRGs) for use in Medicare claims.

4.2 Electronic Sources

All data elements are in defined fields in electronic claims.

4.3 Data Collection Strategy

4.3.1 Data Collection Strategy Difficulties

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.

4.3.1.1 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. As such, there is a trade-off between efficiency (accessing the data in a timely manner) and accuracy (waiting until most claims are finalized) when determining the length of the time (i.e., the “claims run-out” period) after which to pull claims data. To determine the appropriate claims run-out period, 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 and testing purposes. If this measure were used in a CMS program, calculation and reporting would be done in line with that program’s reporting practices.

4.3.1.2 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. 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 adjust accurately 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.

4.3.1.3 Sampling

During measure testing, Acumen noted that episodes in which the beneficiary died prior to the episode end date exhibited different cost distributions compared to other episodes. To avoid this effect’s potential impact on 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.

5.0 Usability and Use

5.1 Use

5.1.1 Current and Planned Use

The measure was developed for potential use in the Merit-based Incentive Program (MIPS), under a contract with CMS.

5.1.2 Feedback on the Measure and Development Process

5.1.2.1 Technical Assistance Provided During Development or Implementation

Development: Field Testing

Acumen and CMS conducted a national field test of 11 episode-based cost measures developed in 2018, including the Renal or Ureteral Stone Surgical Treatment measure, for a 35-day comment period (October 3 to November 5, 2018). We provided field test reports to a sample of clinician groups and clinicians.²² Each report included information for all measures for which the clinician or clinician group was attributed 10 or more episodes. The testing sample was selected to balance coverage and reliability, 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 does not determine case minimums used for any potential program implementation.

- Total testing sample for 11 episode-based cost measures: 20,852 TINs; 127,530 TIN-NPIs
- Testing sample for Renal or Ureteral Stone Surgical Treatment: 1,722 TINs; 4,434 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 Measure Development Process document, a Frequently Asked Questions document, and a Fact Sheet.²³ During field testing, Acumen conducted education and outreach activities including a national webinar, office hours with specialty societies, and Help Desk support.

5.1.2.2 Technical Assistance with Results

Field Testing

During the feedback period, 2,388 field test reports for episode-based cost measures were downloaded by 403 clinician groups (TINs) and 1,985 clinicians (TIN-NPIs). Stakeholder comments from field testing were summarized for the workgroup to consider in recommending refinements to the measures based on the testing data and feedback.

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 contained the following sheets:

²² The field test reports were available for download from the CMS Enterprise Portal: <https://portal.cms.gov/wps/portal/unauthportal/home/>.

²³ The Measure Development Process, Frequently Asked Questions, and Fact Sheet documents are posted on the MACRA Feedback Page: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html>.

- High-level summary results across all episode-based cost measures being field tested
- Results for each measure including cost measure score and breakdown of episode cost compared to the national average and TIN/TIN-NPIs with a similar patient case mix (or risk profile)
- Drill-down detail for each measure, including more detailed information on potential cost drivers in the TIN/TIN-NPI's episodes. For example:
 - Analysis of utilization and cost for the measure by specific service categories (e.g., outpatient evaluation and management services, procedures, and therapy, hospital inpatient services, emergency room services, post-acute services)
 - Breakdown of costs for Physician/Supplier Part B and inpatient claims (e.g., top 5 most billed services and by risk bracket)
- Episode-level table with detailed information for all episodes attributed to the TIN/TIN-NPI across all measures in the report
 - 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.

A mock field test report can be viewed on the CMS MACRA Feedback webpage.²⁴

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 clinician engagement efforts, as well as CMS, Quality Payment Program, and other available listservs. More detail on this outreach can be found in the Field Test Summary Report on the CMS MACRA Feedback webpage.

Acumen and CMS hosted two office hours sessions in October 2018, to provide an overview of field testing to specialty societies, discuss what information their members would be particularly interested in, and answer any questions. Acumen also hosted two office hours sessions with members of Clinical Subcommittees and workgroups to provide an update on development and field testing. Across all four office hours sessions, there were over 100 attendees.

Acumen worked with the Physician Value helpdesk and QPP Service Center to answer stakeholder questions during field testing and continued to answer questions after the feedback period ended.

Acumen and CMS hosted a national field testing webinar on October 9, 2018 to provide an overview of the measures being field tested and the information available for public comment. The webinar consisted of an hour-long presentation, outlining (i) the cost measure development activities, (ii) field testing activities, (iii) how to access and understand the confidential field test reports, and (iv) the contents of the reports. The presentation was followed by a 30-minute Q&A session. Around 85 comments and questions were received via webinar chat and on the phone.

A post-field testing webinar was held on March 27, 2019 to provide an update on the measures following field testing. The webinar consisted of a 60 minute presentation providing an overview of the basics of measure construction, highlighting refinements made after field testing, and

²⁴ CMS, "Episode-based Cost Measures Mock Field Test Report," *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2018-Mock-report-for-Episode-Based-Cost-Measures.xlsx>.

summarizing the testing done on the measures. This presentation was followed by a Q&A session.²⁵

5.1.2.3 Feedback on Measure Performance and Implementation

Field Testing

In total, Acumen received 68 survey responses and 25 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 contained general and detailed questions on the reports themselves, questions on the supplemental documentation, and questions on the measure specifications.

Pre-Rulemaking

CMS received 37 comments on the 11 episode-based cost measures included in the Measures Under Consideration List released in December 2018. This included 3 comments for the Renal or Ureteral Stone Surgical Treatment cost measure. After the Measure Applications Partnership (MAP) Clinician Workgroup meeting in December 2018, there was another public comment period on their preliminary recommendations, which received 7 comments across the 11 measures, with 2 comments specific to the Renal or Ureteral Stone Surgical Treatment cost measure.²⁶ These public comment periods were facilitated by NQF. Stakeholders were able to submit their comments via the NQF website.

5.1.2.4 Feedback from Providers being Measured

Field Testing

The Field Testing Feedback Summary Report 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 remains an important aspect of the measure development process.* Stakeholders expressed appreciation for the opportunity to provide feedback during field testing and for CMS' continued efforts to involve them in the measure development process. Commenters also valued the decision to operationalize previously collected feedback, as demonstrated through the addition of measure-specific workgroups to the development process.
- *Field test reports present useful information for understanding clinician performance, though reduced complexity could encourage more clinician participation.* Stakeholders praised the presentation and content of the field test reports. However, the complexity of the information presented in the reports was a challenge for some stakeholders.
- *Improved supplemental field testing materials are helpful but can be further refined.* Some stakeholders found the supplemental field testing materials to be informative and thorough, providing useful information on field testing and the specifications of the cost measures. However, many noted that although the materials are comprehensive, they remain lengthy and complex, and they believe the amount of information provided is too overwhelming to be useful.
- *Ample time for review of field testing reports and materials is vital to collecting meaningful stakeholder feedback.* Some stakeholders suggested the field testing period be extended or kept open, given the large amount and complexity of the information that was presented.

²⁵ Webinar Recordings, Slides, and Transcripts, *QPP Webinar Library*, <https://qpp.cms.gov/about/webinars>.

²⁶ Measure Applications Partnership, *National Quality Forum*, https://www.qualityforum.org/Setting_Priorities/Partnership/Measure_Applications_Partnership.aspx.

- *Transparent Clinical Subcommittee and measure-specific workgroup selection and voting encourages buy-in from stakeholders.* Some stakeholders expressed concern with the selection and voting processes for the Clinical Subcommittees and workgroups, highlighting that a transparent approach to member selection would ensure an appropriate mix of specialties and clinician types.
- *Field test report access continues to present challenges for stakeholders.* Some stakeholders noted that they faced difficulties creating accounts and downloading their field test reports from the CMS Enterprise Portal and these challenges may have negatively impacted the number of clinicians that were able to participate in field testing. Stakeholders urged CMS to communicate directly with clinicians receiving field test reports and to find an alternative for delivering and accessing the reports.

The report additionally contains measure-specific feedback, which was used as the basis for the post-field testing refinements that were made to the measures, summarized below:

- 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.

5.1.2.5 Feedback from Other Users

Pre-Rulemaking

The MAP recognized the importance of cost measures to the MIPS program and conditionally supported the Renal or Ureteral Stone Surgical Treatment cost measure pending NQF endorsement. Specifically, the MAP encouraged the NQF Cost and Efficiency Standing Committee to 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. The MAP also encouraged the Standing Committee to examine the exclusions in this measure to ensure appropriate attribution.

5.1.2.6 Consideration of Feedback

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 testing commenters and an expert clinician workgroup comprised of subject matter and measure-development experts.

After completing field testing, Acumen compiled the feedback provided through the survey and comment letters into a measure-specific report, which was then provided to the expert clinician workgroup, along with empirical analyses to inform their discussion and evaluation of any refinements needed to ensure that the measure is capturing what it was intended to capture.

The changes to the Renal or Ureteral Stone Surgical Treatment measure made after consideration of field testing analyses and stakeholder feedback are:

- **Episode Windows:** Changed the post-trigger period to 30 days
- **Service Assignment:**

- Removed costs of secondary planned procedure (identified with modifier code 58) occurring in the post-trigger window of an episode triggered from the initial procedure
- Assigned costs for post-acute care services associated with the following diagnoses:
 - N10 Acute pyelonephritis
 - N132 Hydronephrosis with renal and ureteral calculous obstruction N136 Pyonephrosis
 - N200 Calculus of kidney
 - N201 Calculus of ureter
 - N390 Urinary tract infection, site not specified
 - Z48816 Encounter for surgical aftercare following surgery on the genitourinary system
- Used a 7-day post-trigger window for assigning inpatient rehabilitation facility and home health services
- **Exclusions:** Added an exclusion for non-trigger codes for urologic procedures performed on the same day
- **Risk Adjustors:** Removed the following risk adjustors already captured through standard HCC:
 - Acute Kidney Injury
 - Chronic Kidney Disease, Stages 4 and 5
 - ESRD

5.2 Usability

5.2.1 Improvement

n/a. The measures have not yet been implemented, and as such have not had influence over performance.

5.2.2 Unexpected Findings

n/a. There were no unexpected findings during the development and testing of this measure

5.2.3 Unexpected Benefits

n/a. There were no unexpected benefits during the development and testing of this measure.

6.0 Related and Competing Measures

6.1 Relation to Other Cost Measures

There are currently no related NQF-endorsed or non-NQF-endorsed cost measures that address this same measure focus or target population. There are no competing NQF-endorsed or non-endorsed cost measures that address both this same measure focus *and* at this same target population.

6.2 Harmonization

n/a

6.3 Competing Measures

n/a

Contact Information

Measure Steward Point of Contact

Organization: Centers for Medicare & Medicaid Services

Name: Joel Andress

Email Address: joel.andress@cms.hhs.gov

Phone Number: (410) 786-5237

Developer Point of Contact

Organization: Acumen, LLC

Email Address: macra-cost-measures-info@acumenllc.com

Phone Number: (650) 558-8882

Other Additional Information

Renal or Ureteral Stone Surgical Treatment Measure-Specific Workgroup Members:

Adam Weinstein, MD, Renal Physicians Association

Alec Koo, MD, American Urological Association

Brian Matlaga, MD, MPH, American Urological Association

C. Ryan Barnes, MD, American Association of Clinical Urologists

Darryl Zuckerman, MD, Society of Interventional Radiology

Deborah Kaye, MD, MSc, American Urological Association

Jeremy Shelton, MD, MSHS, American Urological Association

Lindsey Herrel, MD, MS, American Urological Association

Parth Modi, MD, American Urological Association

Purushottam Dixit, MD, Society of Interventional Radiology

Robert Dowling, MD, American Urological Association

Tanaz Ferzandi, MD, MBA, MA, American Urogynecologic Society

The Renal or Ureteral Stone Surgical Treatment workgroup is composed from the larger Urologic Disease Management Clinical Subcommittee. The composition list of the Clinical Subcommittee is included in the [Episode-Based Cost Measures Development Process document](#).²⁷

²⁷ CMS, "Episode-Based Cost Measure Field Testing Measure Development Process," *MACRA Feedback Page*, <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/2018-measure-development-process.pdf>.