Quality Payment
PROGRAM

Merit-Based Incentive Payment System (MIPS): Cataract Removal with Intraocular Lens (IOL) Implantation Measure

Measure Justification Form

December 2023

Table of Contents

1.0 Introduction

This Measure Justification Form (MJF) provides results for the testing and evaluation of the Cataract Removal with Intraocular Lens (IOL) Implantation measure. The form is intended to provide detailed information about the testing conducted on this measure, and accompanies the Measure Methodology[1] and Measure Codes List[2] file, which together, comprise the specifications for this cost measure.

1.1 Project Title

Physician Cost Measure and Patient Relationship Codes

1.2 Date

Information included is current on December 8, 2023

1.3 Project 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 Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) requirements. The contract name is "Physician Cost Measure and Patient Relationship Codes (PCMP)." The contract number is 75FCMC18D0015, Task Order 75FCMC19F0004.

1.4 Measure Name

Cataract Removal with Intraocular Lens (IOL) Implantation Episode-Based Cost Measure

1.5 Type of Measure

Cost/Resource Use

1.6 Measure Description

The Cataract Removal with IOL Implantation episode-based cost measure evaluates a clinician's or clinician group's risk-adjusted and specialty-adjusted cost to Medicare for patients who undergo a procedure for the cataract removal with IOL implantation during the performance period. This procedural measure includes the costs of services that are clinically related to the attributed clinician's role in managing care during a Cataract Removal with IOL Implantation episode.

[1]CMS, "Cataract Removal with IOL Implantation Measure Methodology," *Cost Measure Information Page,* https://www.cms.gov/medicare/quality/value-based-programs/cost-measures
[2]CMS, "Cataract Removal with IOL Implantation Measure Codes List" *Cost Measure Information Page*, https://www.cms.gov/medicare/quality/value-based-programs/cost-measures

2.0 Importance

2.1 Evidence to Support the Measure Focus

The Cataract Removal with IOL Implantation measure was developed for use in the Merit-based Incentive Payment System (MIPS) to meet the requirements of the Social Security Act section 1848(r), added by the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). MIPS aims to reward high-value care by measuring clinician performance through four areas: quality, improvement activities, promoting interoperability, and cost. Each category assesses different aspects of care, and the categories are weighted to combine into one composite score. CMS is introducing MIPS Value Pathways (MVPs) to align and connect quality measures, cost measures, and improvement activities across performance categories of MIPS for different specialties or conditions. MVPs aim to provide a holistic assessment of clinician value for a specific type of care to achieve better healthcare outcomes and lower patient costs.

The use of cost measures is required by statute, and their purpose is to assess resource use. To be effective, they should capture costs related to a clinician's care decisions and account for factors outside their influence. This measure provides clinicians with information about their care costs that they can use to understand the costs associated with their decision-making. Clinicians play an important role in the variation in healthcare expenditures due to their ability to affect costs.[3] A cost measure offers an opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better quality of care through changes in clinical practice.

According to the literature and feedback received through stakeholder input activities, this measure's focus represents an area with opportunities for improvement. As discussed in the rest of this section, primary opportunities for improving care and reducing cost outcomes include mitigation of costly complications that require long-term management, mitigation of complications that result in a return to the operating room, and reduce preoperative testing.

The likelihood of developing cataracts increases with age. According to the American Academy of Ophthalmology, 24.4 million Americans aged 40 and older are affected by cataracts, and by age 75, approximately half of Americans have cataracts. With the age of the US population increasing, cataracts will continue to become a greater concern over time. The National Eye Institute estimates the total number of people with cataracts will increase to 39 million by 2030 and 50 million by 2050. The rate of cataract surgery has also increased over the past several decades. From 1980 to 2003, the cataract surgery rate for Medicare beneficiaries increased from 13.4 to 61.8 persons per 1,000 person-years. However, cataracts are reversible with removal surgery, which is highly successful and the most commonly performed surgery in the United States at approximately 3.7 million cases per year. Among Medicare beneficiaries, a total of 14,396,438 cataract surgeries were performed between 2011 and 2019.

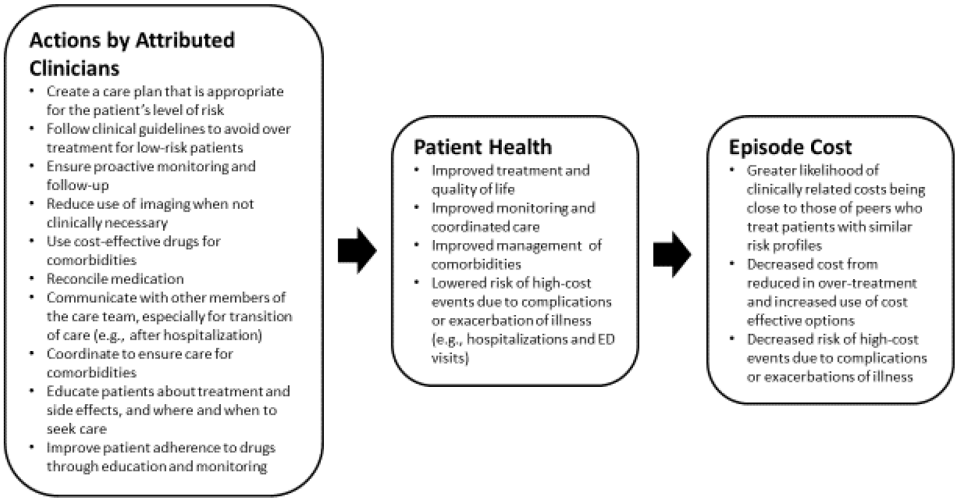
Cataract Removal with IOL Implantation procedures that preserve sight are highly cost effective because vision loss is economically burdensome. It is estimated that approximately 90% of Medicare beneficiaries with coded diagnoses of vision loss incur higher costs than those with normal vision. Considering the entire Medicare population, blindness and vision loss cost $2.14

[3]David Cutler et al., "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," American Economic Journal: Economic Policy 11, no. 1 (February 1, 2019): 192–221, https://doi.org/10.1257/pol.20150421.

billion in 2003. Furthermore, Medicare beneficiaries with vision loss tend to have higher medical expenditures than beneficiaries without vision loss.

2.1.1 Logic Model

Figure 1: Logic Model of Steps between Actions by Attributed Clinicians and Episode Cost



2.2 Performance Gap

2.2.1 Rationale

The Cataract Removal with Intraocular Lens (IOL) Implantation episode-based cost measure assesses costs related to cataract surgery, addressing a current measurement gap in the MIPS cost performance category. Furthermore, an environmental scan of the literature identified four critical areas for improving care and reducing costs, including:

1. Mitigation of costly complications that require long-term management
2. Mitigation of complications that result in a return to the operating room
3. Reduction of preoperative testing

The American Academy for Ophthalmology (AAO) preferred practice pattern recommends that surgeons should be aware of and prepared to manage high-risk characteristics that may complicate cataract surgery and lead to long-term management. The most significant complication to be treated is endophthalmitis—between 2010-2014, approximately 1.2 per 1000 cataract surgery cases developed endophthalmitis, and each patient incurs substantial treatment costs as a result. Endophthalmitis incidence rates are typically higher in combined cataract procedures compared to standalone cataract surgery. Among Medicare beneficiaries who underwent cataract surgery between 2011 and 2019, the 90-day postoperative endophthalmitis rate was 1.36 versus 1.3 per 1,000 for combined cataract and standalone cataract procedures, respectively.

IOL complications resulting in reoperation are rare but can vary depending on the design and material of a particular IOL, the surgical technique, and other risk factors. These complications include dislocation, decentration, retained lens fragments, and retinal detachment. Dislocation and decentration have been reported with virtually all IOL materials and models, including both

one- and three-piece designs. Additionally, patients with an axial length of more than 25 millimeters, younger patients, and males were more at risk of retinal detachment.

Preoperative testing does not reduce the risk of adverse outcomes or improve outcomes for cataract surgery patients. Yet, an observational study of Medicare beneficiaries undergoing cataract surgery in 2011 found that 53% had at least one preoperative test the month before surgery, resulting in higher expenditures on testing and office visits, $4.8 million and $12.4 million, respectively. An updated analysis from 2018 found that the mean number of tests per month per patient in the 30 days preceding cataract removal surgery was 1.8 tests, with a total estimated cost of $22.7 million, and an estimated routine preoperative testing cost to Medicare of up to $45.4 million annually.

The Cataract Removal with Intraocular Lens (IOL) Implantation episode-based cost measure was recommended for development through feedback gathered during a public comment period. The public recommended this measure because of its impact in terms of patient population and clinician coverage, and the opportunity for incentivizing cost-effective, high-quality clinical care in this clinical area. A measure-specific Clinician Expert Workgroup was then convened with clinicians, healthcare experts, and patient representatives who have appropriate experience to provide extensive, detailed input on this measure throughout its development.

Based on input from a Technical Expert Panel (TEP) meeting and Clinical Expert Workgroup, as well as internal analyses on the measure's performance, we recommend the revised Cataract Removal with Intraocular Lens (IOL) Implantation measure for implementation in the MIPS cost performance category for the CY 2025 performance period.

2.2.2 Performance Scores

Table 1 displays the distribution of the revised measure scores for clinicians and clinician groups identified by a Tax Identification Number (TIN) and individual clinicians identified by a combination of a Tax Identification Number and National Provider Identifier (TIN-NPI). These results align with expectations based on our review of the literature and demonstrate that a performance gap in cost measure performance at both the clinician and clinician group levels.

There are variations in cost performance in the measure score for both TIN and TIN-NPI, as evidenced by the interquartile range, greater than 10%. Additionally, the 90th percentile score was more than 70% higher than the 10th percentile score for both TIN and TIN-NPI levels. The variation in the measure score, indicated by the interquartile range and standard deviation, is in the thousands of dollars, highlighting an opportunity for improvement in the costs of care for Cataract Removal with Intraocular Lens (IOL) Implantation by closing the gap between the most and least efficient providers.

Table 1. Distribution of the Measure Score

Metric	TIN	TIN-NPI
Count	4,080	8,724
Mean Score	$3,172	$3,155
Score Standard Deviation	$394	$395
Minimum Score	$1,227	$775
Maximum Score	$5,070	$5,070
Score Interquartile Range (IQR)	$341	$340
Score Percentile		

Metric	TIN	TIN-NPI
10th	$2,858	$2,837
20th	$2,944	$2,930
30th	$2,993	$2,978
40th	$3,038	$3,025
50th	$3,093	$3,077
60th	$3,175	$3,153
70th	$3,254	$3,242
80th	$3,373	$3,360
90th	$3,623	$3,605

2.2.3 Disparities

Data on how the measure, as specified, addresses disparities is described in Sections 3.1.7 and 3.5.5.

3.0 Scientific Acceptability

3.1 Data Sample Description

Testing is based on the full population of measured entities and patients meeting inclusion and exclusion criteria for the measure, not based on a sample.

3.1.1 Type of Data Used for Testing

Medicare administrative claims data from the Common Working File (CWF), Long-Term Care Minimum Data Set (LTC MDS), and Medicare Enrollment Database (EDB).

3.1.2 Specific Dataset Used for Testing

The Cataract Removal with Intraocular Lens (IOL) Implantation measure uses Medicare Part A and Part B claims data maintained by CMS. Part A and B claims data are used to construct episodes of care, calculate episode costs, and develop risk adjustors. Episode costs undergo payment standardization and risk adjustment to ensure accurate comparisons of costs across clinicians. Payment standardization adjusts the allowed amount for a Medicare service to limit observed differences in costs to those that may result from healthcare delivery choices.

Data from the EDB are utilized to determine beneficiary-level exclusions and secondary risk adjustors, specifically Medicare Parts A, B, and C enrollment, primary payer, disability status, end-stage renal disease (ESRD), patient birth dates, and patient death dates. The risk adjustment model also accounts for expected differences in payment for services provided to patients in long-term care based on data from the MDS. Specifically, the MDS is used to create the long-term care indicator variable in risk adjustment.

3.1.3 Dates of the Data Used in Testing

Cataract Removal with Intraocular Lens (IOL) Implantation episodes ending from January 1, 2022, through December 31, 2022.

3.1.4 Levels of Analysis Tested

The measure was tested at group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.1.5 Entities Included in the Testing and Analysis

Table 2 shows the individual clinician (identified by combination of TIN and NPI) and clinician group/practice (identified by TIN) included in the testing of the Cataract Removal with Intraocular Lens (IOL) Implantation measure.

Table 2: Measured Entities Demographics

Metric	TIN		TIN-NPI	
	Count	%	Count	%
Count	4,080	100.00%	8,724	100.00%
Number of Episodes Attributed	-	-	-	-
10-19 Episodes	376	9.22%	1,064	12.20%
20-39 Episodes	657	16.10%	1,790	20.52%
40-59 Episodes	470	11.52%	1,360	15.59%
60-79 Episodes	354	8.68%	1,028	11.78%
80-99 Episodes	295	7.23%	795	9.11%
100-199 Episodes	863	21.15%	1,918	21.99%
200-299 Episodes	368	9.02%	493	5.65%

Metric	TIN		TIN-NPI	
	Count	%	Count	%
300+ Episodes	697	17.08%	276	3.16%
Census Region	-	-	-	-
Northeast	895	21.94%	1,799	20.62%
Midwest	800	19.61%	1,879	21.54%
South	1,404	34.41%	3,067	35.16%
West	945	23.16%	1,937	22.20%
Unknown	36	0.88%	42	0.48%

3.1.6 Patient Cohort Included in the Testing and Analysis

Table 3 shows the patient population for the Cataract Removal with Intraocular Lens (IOL) Implantation measure testing. It comprises of Medicare beneficiaries enrolled in Medicare Parts A and B who undergo a procedure for cataract removal with IOL implantation, triggering a Cataract Removal with Intraocular Lens (IOL) Implantation episode.

Table 3: Beneficiary Demographics

Metric	Value
Count	800,983
Mean Age	74.42 years
Female %	60.90%

3.1.7 Social Risk Factors Included in Analysis

The analysis of social risk factors (SRFs) focused on examining the impact of Dual Medicare and Medicaid enrollment status on the measure. Table 4 outlines variables that may indicate SRFs and their advantages and disadvantages as indicators of individual-level SRFs. On balance, the analysis used dual Medicare and Medicaid enrollment status as the proxy of SRFs due to its broad availability in claims data, accurate measurement at the individual level, and wide acceptance as a powerful indicator of health outcomes.[4]

Table 4: Social Risk Factors Available for Analysis

Variable	Advantages	Disadvantages	Used in Testing
Dual Medicare and Medicaid enrollment status	• Available for all beneficiaries • Most powerful predictor of poor outcomes[4]	• Variation in Medicaid eligibility across states	Yes
Race/Ethnicity	• Available for most beneficiaries, except for ambiguous categories of "Unknown" or "Other"	• Social risk driven by someone's race is often correlated with and partially captured by dual status[4]	No

[4] Office of the Assistant Secretary for Planning and Evaluation. "Second report to Congress on social risk and Medicare's value-based purchasing programs." (2020) https://aspe.hhs.gov/pdf-report/second-impact-report-to-congress

Variable	Advantages	Disadvantages	Used in Testing
		• Only 5 categories available, which may lack granularity to fully capture disparities[5,6]	
ICD-10 Z codes for social determinants of health	• Reflects individual-level factors that influence health status and contact with health services	• Not routinely and consistently coded on claims, only available for 0.1% of all fee-for-service claims in 2019[7]	No
American Community Survey	• Can link beneficiary's zip code to socioeconomic (SES) measurement of their neighborhood • Many SES indices can be derived from the survey data (e.g., AHRQ index, deprivation index)	• Only a proxy measure, not always accurate at individual-level	No

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, group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.2.2 Method of Reliability Testing

Data Element Reliability

The Cataract Removal with Intraocular Lens (IOL) Implantation 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.

- First, CMS routinely conducts data analyses to identify potential problem areas and detect fraud and audits necessary data fields used in this measure, including diagnosis and procedure codes and other elements consequential to payment. Specifically, CMS works with Zone Program Integrity Contractors, formerly Program Safeguard Contractors, to ensure program integrity; the agency also uses Recovery Audit Contractors to identify and correct for underpayments and overpayments.
- Second, CMS also uses the Comprehensive Error Rate Testing (CERT) Program to ensure that Medicare payments are correct under coverage, coding, and billing rules. CMS continues to perform corrective actions and gives providers additional education to ensure accurate billing.

[5] Nguyen, Kevin H., Kaitlyn P. Lew, and Amal N. Trivedi. "Trends in Collection of Disaggregated Asian American, Native Hawaiian, and Pacific Islander Data: Opportunities in Federal Health Surveys." *American Journal of Public Health* (2022).
[6] Kader, Farah, Lan N. Doan, Matthew Lee, Matthew K. Chin, Simona C. Kwon, and Stella S. Yi. "Disaggregating Race/Ethnicity Data Categories: Criticisms, Dangers, And Opposing Viewpoints", *Health Affairs Forefront* (2022).
[7] Centers for Medicare and Medicaid, Office of Minority Health. "Utilization of Z Codes for Social Determinants of Health among Medicare Fee-for-Service Beneficiaries." (2019) https://www.cms.gov/files/document/z-codes-data-highlight.pdf

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

Clinician-level 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 how much of the variation in the measure score is explained by differences among clinicians' performance (i.e., signal) rather than random variation (i.e., statistical noise) among 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, suggesting that the measure is effectively capturing the systematic differences between the clinician and their peer cohort.

3.2.3 Statistical Results from Reliability Testing

Data Element Reliability

Between 2005 and 2020, CMS Comprehensive Error Rate Testing (CERT) estimates that proper payment, which includes payments that met Medicare coverage, coding, and billing rules, ranged from 87.3% to 93.7% of total payments each year.[8] The fiscal year 2022 Medicare fee-for-service program proper payment rate was 92.5%.[9]

Clinician-level Reliability

The table below shows reliability metrics at the 20-episode testing volume threshold. While higher thresholds generally yield higher reliability results, these increases must be considered against decreasing the number of clinicians and clinician groups eligible for the measure, which would limit the applicability of measures to larger group practices and potentially limit the impact of the measure in encouraging performance improvement. For testing purposes, we used a 20-episode volume threshold. If the measure is implemented in MIPS in the future, CMS will establish a case minimum through notice-and-comment rulemaking.

[8]Comprehensive Error Rate Testing (CERT) Program. "Appendices Medicare Fee-for-Service 2020 Improper Payments Report". Table A6. https://www.cms.gov/files/document/2020-medicare-fee-service-supplemental-improper-payment-data.pdf-1.
[9]Ibid.

Table 5: Reliability at the Accountability Entity Level

Reporting Level	Entities Meeting Case Minimum	Mean Reliability	Median Reliability	% Above 0.4	% Above 0.7
TIN	3,704	0.98	0.98	100.00%	100.00%
TIN-NPI	7,660	0.97	0.98	100.00%	100.00%

3.2.4 Interpretation

The results of the data element testing show very high reliability of the critical data elements used by the measure. At the entity level, the measure is highly reliable for both the TIN and TIN-NPI reporting levels with a mean reliability of 0.98 and 0.97, respectively. A measure with high reliability suggests that performance comparisons across clinicians reflect systematic differences in actual performance better. Based on existing scientific evidence regarding different interpretations and methods of estimating reliability, CMS finalized in the CY 2022 Physician Fee Schedule (86 FR 64996) rule that the 0.4 threshold for mean reliability continues to be appropriate for indicating moderate reliability for performance measures in the Cost category in the MIPS program. Mean reliability levels above 0.7 continue to demonstrate high reliability for cost measures, as previously established in the CY 2017 Quality Payment Program final rule (81 FR 77169 through 77171).[10] Additionally, at each testing volume threshold, 100% of TINs and TIN-NPIs meet or exceed the moderate reliability threshold of 0.4 and 100% are above the high reliability threshold of 0.7.

3.3 Validity Testing

3.3.1 Level of Validity Testing

The validity of the measure was tested using empirical validity at the accountable entity level (TIN and TIN-NPI).

3.3.2 Method of Validity Testing

Face Validity

The Cataract Removal with Intraocular Lens (IOL) Implantation measure was developed through a structured, iterative process for gathering detailed input on the measure from recognized clinician experts. Experts in this clinical area evaluated specifications 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 between good and poor performance).

In developing this measure, Acumen incorporated input from:

(i) a Cataract Removal with Intraocular Lens (IOL) Implantation Clinician Expert Workgroup;
(ii) a Technical Expert Panel (TEP); and
(iii) the Person and Family Partners.

[10] CMS, "Medicare Program; CY 2022 Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Provider Enrollment Regulation Updates; and Provider and Supplier Prepayment and Post-Payment Medical Review Requirements," 86 FR 64996-66031.

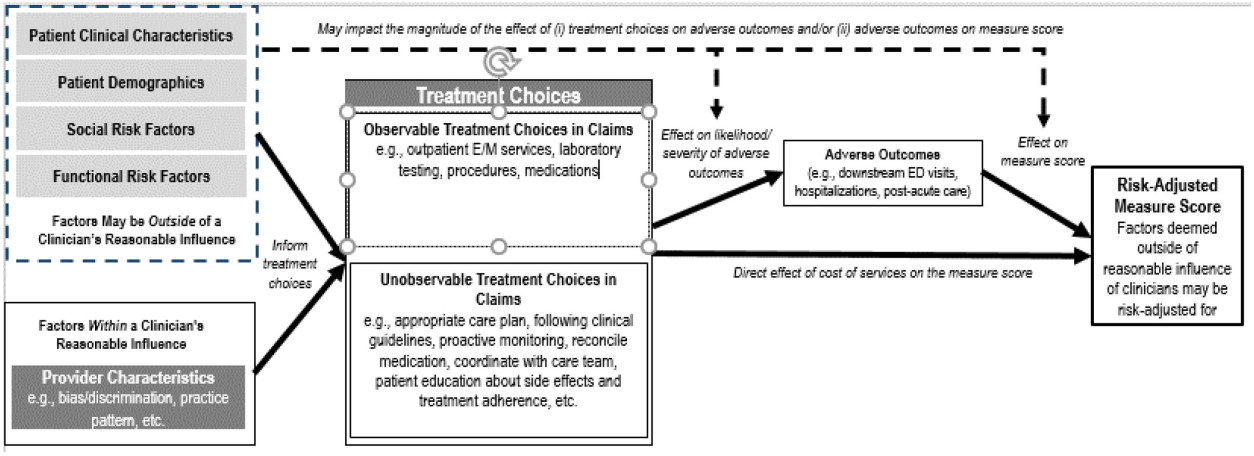
This process is detailed in the Episode-Based Cost Measures Development Process document posted on the Cost Measures Information Page.[11]

One of the primary roles of the Clinician Expert Workgroup is to develop service assignment rules for the cost measure. These service assignment rules seek to ensure clinicians are evaluated on services and costs that are clinically related to the attributed clinician's role in managing care during each episode from 60 days prior to the clinical event that opens, or "triggers," the episode through 90 days after the trigger, thus limiting cost variation unrelated to clinician care in this measure. Therefore, assigned services are services that the Clinical Expert Workgroup believed an attributed clinician could influence their occurrence, frequency, or intensity.

Empirical Validity Testing

Validity is a criterion used to assess whether the cost measure can quantify the construct it aims to measure, which is the cost directly related to treatment choices and the cost of adverse outcomes resulting from care. We evaluated the empirical validity of the Cataract Removal with Intraocular Lens (IOL) Implantation measure by estimating the effect of relevant treatment choices on the measure score using multiple regression, based on the conceptual model outlined in Figure 2.

Figure 2: Conceptual Model of Treatment Choices on the Measure Score



The cost measure is designed to reflect costs directly related to treatment choices and the cost of adverse outcomes resulting from care. Therefore, treatment choices, whether observable in claims or otherwise, by an attributed clinician can directly impact the measure score or indirectly when they are mediated through the cost of adverse outcomes. In turn, the cost of adverse effects contributes to the total cost captured by the measure score.

This analysis first estimates the association between treatment choices and the measure score while controlling for the cost of adverse outcomes to demonstrate that the score reflects both the direct and indirect effects of treatment choices. Then, the association between treatment choices and the cost of adverse outcomes is estimated to illustrate the indirect effect.

[12]CMS, Cost Measures Information Page, https://www.cms.gov/medicare/quality/value-based-programs/cost-measures.

Generally, adverse outcomes include non-trigger inpatient hospitalizations, non-trigger emergency room visits, and post-acute care. The remaining cost categories are generally considered treatment. For each of these categories, the regression models use the mean cost across episodes that were attributed to an individual clinician. The measure score is represented by a clinician's mean observed cost over expected cost ratio across their attributed episodes.

3.3.3 Statistical Results from Validity Testing

Empirical Validity Testing

Table 6 shows two regression models for each reporting level. Model 1 illustrates the effect on the clinicians' mean observed cost to expected cost ratio for each additional one thousand dollars of a cost category assigned to an episode, on average, while holding the remaining categories of cost constant. Model 2 demonstrates the effect on the mean cost of adverse events for each additional one thousand dollars of a cost category assigned to an episode, on average, while holding the remaining categories of cost constant.

Table 6. Estimated Effect on Treatment Choices on the Measure Score

Service Categories	Coefficient in Thousands [95% Confidence Interval] (p-value)			
	TIN		TIN-NPI	
	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices
Adverse Events	0.06 [-0.89,1.02] (p = 0.90)	-	-0.49 [-1.07,0.09] (p = 0.10)	-
Outpatient Evaluation & Management Services	0.18 [0.14,0.22] (p < 0.01)	0.00 [0.00,0.00] (p = 0.33)	0.21 [0.19,0.24] (p < 0.01)	0.00 [0.00,0.00] (p = 0.53)
Major Procedures	0.16 [0.07,0.25] (p < 0.01)	0.00 [0.00,0.00] (p = 0.70)	0.11 [0.04,0.18] (p < 0.01)	0.00 [0.00,0.00] (p = 0.22)
Ambulatory/Minor Procedures	0.06 [0.06,0.07] (p < 0.01)	0.00 [0.00,0.00] (p = 0.07)	0.07 [0.06,0.07] (p < 0.01)	0.00 [0.00,0.00] (p = 0.04)
Laboratory, Pathology, and Other Tests	0.30 [-0.10,0.70] (p = 0.15)	0.00 [-0.01,0.01] (p = 0.87)	0.48 [0.20,0.77] (p < 0.01)	0.01 [0.00,0.02] (p = 0.09)
Imaging Services	-0.02 [-0.14,0.11] (p = 0.81)	0.00 [-0.01,0.00] (p = 0.25)	-0.10 [-0.18,-0.01] (p = 0.03)	0.00 [-0.01,0.00] (p < 0.01)
Durable Medical Equipment and Supplies	-0.08 [-0.24,0.08] (p = 0.31)	0.00 [-0.01,0.00] (p = 0.60)	-0.13 [-0.24,-0.02] (p = 0.02)	0.00 [-0.01,0.00] (p = 0.09)

Chemotherapy and Other Part B-Covered Drugs	0.27 [0.25,0.29] (p < 0.01)	0.00 [0.00,0.00] (p = 0.12)	0.28 [0.27,0.29] (p < 0.01)	0.00 [0.00,0.00] (p = 0.12)
Anesthesia Services	0.15 [0.08,0.22] (p < 0.01)	0.00 [0.00,0.00] (p = 0.61)	0.17 [0.13,0.21] (p < 0.01)	0.00 [0.00,0.00] (p = 0.68)
All Other Services Not Otherwise Classified	0.04 [-0.09,0.18] (p = 0.54)	0.00 [-0.01,0.00] (p = 0.60)	0.08 [-0.02,0.17] (p = 0.11)	0.00 [0.00,0.00] (p = 0.38)

3.3.4 Interpretation

Validity is a criterion used to assess whether the cost measure can quantify the construct it aims to measure, which is the cost directly related to treatment choices and the cost of adverse outcomes resulting from care. Validity is evaluated empirically by estimating the effect of relevant treatment choices on the measure score. This analysis first estimates the correlation between treatment choices and the measure score while controlling for adverse outcomes.

The cost measure is designed to reflect the cost directly related to treatment choices, as well as the cost of adverse outcomes as a result of care. Therefore, treatment choices, either observable in claims or otherwise, by an attributed clinician can directly impact the measure score or indirectly when they're mediated through the cost of adverse outcomes. The cost of adverse outcomes, in turn, contributes to the total costs that are captured by the measure score.

To demonstrate that the measure score is reflective of both the direct and indirect effects of treatment choices, this analysis first estimates the association between treatment choices and the measure score while controlling for the cost of adverse outcomes. Then, the association between treatment choices and the cost of adverse outcomes is estimated to demonstrate the indirect effect.

Generally, adverse outcomes are non-trigger inpatient hospitalizations, non-trigger emergency room visits, and post-acute care. The remaining service categories are generally considered treatment. For each of these categories, the regression models use the mean cost across episodes that were attributed to an individual clinician. The measure score is represented by a clinician's mean observed cost over expected cost ratio across their attributed episodes

Overall, the results demonstrate that the cost measure is reflective of both the cost directly related to treatment choices, as well as the cost of adverse outcomes as a result of care (Table 6). Therefore, there is evidence that the measure is capturing what is supposed to measure.

Model 1 shows that there is not a statistically significant relationship between costs of adverse events and the measure score. The cost for outpatient evaluation and management services is associated with slightly worse measure score for both TIN and TIN-NPIs. Additionally, the costs of major or minor procedures and chemotherapy and other Part B drugs are associated with worse measure scores. Model 2 shows that none of these services are associated with the cost of adverse events. The results suggest that these may be potential areas of overuse and reducing the use of these services when appropriate could result in improved cost measure performance. Exclusions Analysis

3.3.5 Method of Testing Exclusions

Exclusions are used in the Cataract Removal with Intraocular Lens (IOL) Implantation measure to ensure a comparable patient population within the scope of the measure's focus on patients who undergo a procedure for cataract removal with IOL implantation and that episodes provide

meaningful information to attributed clinicians. Exclusions are also used as part of data processing so that sufficient data are available to accurately determine episode spending and calculate risk adjustment for each episode.

For the exclusions analysis discussed in this section, we focused on exclusion criteria intended to ensure a comparable patient population.

- Episodes where the patient death date occurred before the episode end date
    - These episodes were excluded as they may not accurately reflect a clinician's performance as the truncated episode window does not capture the full length of care intended by the measure.
- Measure-specific exclusions for patients with 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
    - These episodes are excluded because significant ocular conditions can impact the surgical complication rate/visual outcomes.

Given the rationales for these exclusions, we expect these excluded episodes to have a different 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 each exclusion, we examined the number of episodes and beneficiaries affected, as well as the distributions of observed cost. We then compared the cost characteristics of the excluded episodes to those of episodes included in the measure calculation to assess the distinctness between the two patient cohorts. A full list of the exclusions used for the Cataract Removal with Intraocular Lens (IOL) Implantation measure is provided in the Measure Codes List available on the Cost Measures Information Page.[12]

3.3.6 Statistical Results from Testing Exclusions

Table 7 below presents descriptive statistics of all episodes meeting the measure's triggering logic, excluded episodes, and final reportable episodes at both TIN and TIN-NPI levels. These exclusion criteria ensure that the reportable episode populations are more homogenous and comparable than all episodes meeting triggering logic.

[12]CMS, Cost Measures Information Page, https://www.cms.gov/medicare/quality/value-based-programs/cost-measures.

Table 7: Cost Statistics for Measure Exclusions

Exclusion	Episodes		Observed Cost					
			Mean	Percentile				
	#	% of All Episodes Meeting Triggering Logic		10th	25th	50th	75th	90th
All Episodes Meeting Triggering Logic	912,754	100.00%	$3,179	$1,768	$2,010	$3,137	$3,772	$5,276
Beneficiary Death in Episode	3,659	0.40%	$2,927	$1,712	$1,907	$2,802	$3,626	$4,606
Outlier	16,018	1.75%	$3,677	$566	$838	$2,911	$6,025	$7,406
No Attributed TIN-NPI	22	0.00%	$2,085	$808	$1,647	$1,852	$2,438	$3,374
TIN does not Meet Case Minimum	5,906	0.65%	$2,762	$910	$1,750	$2,474	$3,561	$5,104
TIN-NPI does not Meet Case Minimum	15,469	1.69%	$2,907	$1,049	$1,875	$2,818	$3,630	$5,054
Not in OP, IP, or ASC Setting	173	0.02%	$2,771	$1,491	$1,834	$2,071	$3,267	$5,688
Patients with Significant Ocular Conditions Impacting Surgical Complication Rate/Visual Outcomes	103,554	11.35%	$3,201	$1,732	$1,999	$3,022	$3,793	$5,323
Episode Group-specific Exclusions - Undefined Subgroup	5,660	0.62%	$2,629	$753	$1,200	$2,399	$3,533	$5,607
Reportable Episodes (if all clinicians reported as TIN at the Testing Volume Threshold)	782,366	85.71%	$3,172	$1,793	$2,023	$3,171	$3,762	$4,935
Reportable Episodes (if all clinicians reported as TIN-NPI at the Testing Volume Threshold)	776,792	85.10%	$3,174	$1,794	$2,024	$3,180	$3,763	$4,941

3.3.7 Interpretation

The statistical results present descriptive statistics for all episodes meeting the revised measure's triggering logic, excluded episodes, and final reportable episodes at both TIN and TIN-NPI levels. This supports the exclusion of these episodes to ensure a comparable patient cohort that will yield a clinically coherent measure and meaningful information to attributed clinicians. Further discussion of the results for exclusions applied based on the clinical validity of the study population is provided below.

Overall, exclusion criteria decrease the distribution of observed costs of all episodes meeting trigger logic, as shown by the comparisons between the 10th percentile and 90th percentile.

Episodes classified as outlier cases are excluded because they deviate substantially from the projected cost for a given patient risk profile. Outlier episodes have a much higher mean

observed episode cost of $3,677 compared to $3,179 for all episodes meeting triggering logic. The wide variability of observed episode costs for outlier cases also supports their exclusion. At the 90th percentile the observed cost for outlier episodes is $7,406, compared to $5,276 for all episodes meeting triggering logic.

Based on input from the clinical expert workgroup, episodes of patients with significant ocular conditions impacting surgical complication rate/visual outcomes are excluded because these episodes can be clinically distinct from the overall cataract removal population. These episodes have a slightly higher mean observed cost than all episodes meeting triggering logic, at $3,201.

Episodes where a beneficiary died before the episode end date are excluded because they do not provide sufficient data in the episode window period. These episodes have an almost similar mean observed cost with all episodes meeting triggering logic, at $2,927, likely because the costs are distributed over fewer days than a typical episode.

Episodes that could not be classified into measure subgroups are excluded because they do not include sufficient data about the procedure setting or laterality (i.e., unilateral or bilateral). These episodes also have a similar mean observed cost with all episodes meeting triggering logic, at $2,629, likely because the costs are a composite amount of the four subgroups unclassified. However, these episodes are not included due to a lack of sufficient information necessary for inclusion in the stratification and risk adjustment methodology. The omission of these episodes does not compromise or affect the accuracy and reliability of the measure because the decision for this exclusion is due to the similarity in costs and the low episode count associated with the measure. This measure only includes episodes for procedures performed in HOPDs or ASCs; episodes that occur in physician offices or other settings are excluded to improve comparability of episode costs. .

3.4 Risk Adjustment or Stratification

3.4.1 Method of Controlling for Differences

Differences in case mix are controlled for using a statistical risk model with 107 risk factors and stratification by 4 risk categories.

The risk adjustment model for the Cataract Removal with Intraocular Lens (IOL) Implantation measure adjusts for comorbidities based on the CMS Hierarchical Condition Category (HCC) model, count of HCCs, end-stage renal disease (ESRD) status, disability status, number and types of clinician specialties from which the patient has received care, recent use of institutional long-term care, age, and dual eligibility status.

The model also includes measure-specific factors:

- New or established patients
- Patients with ocular conditions impacting case complexity
- Surgeries performed by a resident under the direction of a teaching physician.

A separate linear regression is run for each sub-group and Medicare Part D enrollment status combination to ensure fair comparison:

- ASC / Bilateral
- ASC / Unilateral
- HOPD / Bilateral
- HOPD / Unilateral

The episode's scaled (i.e., annualized) observed costs are winsorized at the 98th percentile before the regression for each model to handle extreme observations. Full details of the risk adjustment model are in the Measure Codes List File available on the Cost Measures Information Page.[13]

3.4.2 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. 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). Because the CMS-HCC model has already been extensively tested, we focus our testing on the adaptation of the CMS-HCC model to the Cataract Removal with Intraocular Lens (IOL) Implantation measure's patient population.

The workgroup provided input on measure-specific risk adjustors 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 the reasonable influence of the clinician, or any other factors that would help prevent unintended consequences. These additional risk adjustors are listed in the section above.

As previously noted, the risk adjustment model is run on episodes stratified into episode sub-groups, which may qualify as "ordering" of risk factors. Episode sub-groups were also determined based on 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.

3.4.3 Conceptual Model of the Impact of Social Risks

Figure 3 shows the conceptual model that outlines how SRFs can influence the measure score, which is informed by published external research and Acumen's data analysis.[4,14,15,16,17] The conceptual model outlines risk factors that are either known by the literature or informed by the Clinical Expert Workgroup to be within or outside the influence of the attributed clinician. Risk factors, including SRFs, can influence the treatment choices and impact the size of the effect of treatment choices on mitigating the risk and cost of adverse outcomes.

A systematic approach then guides the decision of which factors to include in the risk adjustment model:

1. First, we reviewed the literature to gather known risk factors and drivers of resource use. These factors are usually diagnoses. Therefore, the first set of risk adjustors are commonly the HCCs.
2. Then, we consulted our clinical expert panels on additional factors that are known to be associated with resource use. Together with our clinical expert panel, we reviewed the

[13]CMS, Cost Measures Information Page, https://www.cms.gov/medicare/quality/value-based-programs/cost-measures.
[14]Assistant Secretary of Health and Human Services for Planning and Evaluation. Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs. Washington, D.C. December 2016.
[15]Chen LM, Epstein AM, Orav EJ, Filice CE, Samson LW, Joynt Maddox KE. Association of Practice-Level Social and Medical Risk With Performance in the Medicare Physician Value-Based Payment Modifier Program. JAMA. 2017;318(5):453-461
[16]Medicare Payment Advisory Commission. Beneficiaries Dually Eligible for Medicare and Medicaid. 2018; https://www.macpac.gov/publication/data-book-beneficiaries-dually-eligible-for-medicare-and-medicaid-3/.
[17] Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Second Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program. 2020. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-basedpurchasing-programs

stratified results on episode cost across many patient characteristics. We arrived at the final list of risk adjustors based on those discussions and consensus among the clinical experts.

3. During our testing phases, we also follow a structured and systematic approach to deciding whether SRFs should be adjusted for, further described in Section 3.5.5.

3.4.4 Statistical Results

The literature has extensively tested using the HCC model for Medicare claims data. Although the variables in the HCC model were selected to predict annual cost, CMS has also used this risk adjustment model in several other settings (e.g., Accountable Care Organizations, previous physician Quality and Resource Use Report programs, and other administrative claims-based measures such as the Knee Arthroplasty episode-based cost measure, Total Per Capita Cost (TPCC) cost measure, Medicare Spending Per Beneficiary (MSPB)-PAC cost measure and 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 V24 model can be found in the Evaluation of the CMS-HCC Risk-Adjustment Model report[18] and the Report to Congress: Risk Adjustment in Medicare Advantage[19]. 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 adjustors and sub-groups.

3.4.5 Analyses and Interpretation in Selection of Social Risk Factors

To determine whether it is appropriate to risk adjust for SRFs, the following criteria are considered:

(i) whether there is an association between social risk and performance by examining the coefficient of patient-level dual status when added into the risk model,

(ii) whether the observed association is most influenced by patient-level factors or clinician-level factors by examining the stability of the patient-level dual status coefficient after adding clinician's dual share variable, as well as including clinician's fixed effects,

(iii) whether patient's need or complexity rather than poor quality is driving the observed performance differences by examining the differences in performance on dual patients versus non-dual patients and if there are many clinicians who are able to perform similarly or better on their dual patients than their non-dual patients, and

(iv) the impact of risk adjusting for SRFs by examining the performance shift of clinicians compared to a risk adjustment model that does not risk adjust for SRFs.

Table 8: Coefficient of Patient-level Dual Status under Different Models

Reporting Level	Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
			Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
TIN	ASC / Bilateral	5.35%	-$70.18 (p: <.0001)	-$96.51 (p: <.0001)	-$85.34 (p: <.0001)
	ASC / Unilateral	8.86%	-$37.24 (p: <0.001)	-$59.43 (p: <.0001)	-$59.66 (p: <.0001)

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

[19]CMS, "Report to Congress: Risk Adjustment in Medicare Advantage," https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf.

Reporting Level	Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
			Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
	HOPD / Bilateral	8.05%	-$96.49 (p: <0.001)	-$53.69 (p: .00)	-$42.17 (p: .00)
	HOPD / Unilateral	12.93%	-$70.67 (p: <0.001)	-$35.69 (p: <.0001)	-$28.37 (SD: $8, p: .00)
TIN-NPI	ASC / Bilateral	5.35%	-$70.12 (p: <.0001)	-$95.23 (p: <.0001)	-$85.77 (p: <.0001)
	ASC / Unilateral	8.86%	-$37.24 (p: <.0001)	-$57.40 (p: <.0001)	-$56.69 (p: <.0001)
	HOPD / Bilateral	8.05%	-$96.62 (p: <.0001)	-$55.38 (p: 0.00)	-$32.56 (p: 0.01)
	HOPD / Unilateral	12.93%	-$70.69 (p: <.0001)	-$41.05 (p: <.0001)	-$32.22 (SD: $8, p: <.0001)

Table 9: Mean Ratio of Episode Observed Cost to Expected Cost (O/E) Stratified by Clinician's Dual Share and Patient's Dual Status

Dual Share	TIN			TIN-NPI		
	All Episode	Dual Episodes	Non-Dual Episodes	All Episodes	Dual Episodes	Non-Dual Episodes
ALL	1.00	0.99	1.00	1.00	0.98	1.00
0%-20%	1.00	-	1.00	1.00	-	1.00
21%-40%	1.01	0.98	1.01	1.00	0.98	1.00
41%-60%	1.01	0.99	1.01	1.00	0.98	1.00
61%-80%	1.01	0.99	1.01	1.00	0.99	1.00
81%-100%	1.00	0.99	1.00	1.00	0.99	1.00

Table 10. Proportions of Clinicians Who Perform Significantly Worst, Equally Well, or Significantly Better on Their Dual Episodes than Non-Dual Episodes

Reporting Level	Significantly Worse	Equally Well	Significantly Better
TIN	3.71%	86.42%	9.86%
TIN-NPI	3.57%	86.88%	9.55%

Table 11. Clinicians' Performance Shift after Adding a Dual Status Risk Adjustor

TIN or TIN-NPI	Proportion of Clinicians Affected at Various Levels of Performance Shift	
	Ranking Shift by 1% or more	Ranking Shift by 5% or more
TIN	35.87%	1.16%
TIN-NPI	34.82%	1.09%

The results suggest that there is a statistically significant negative association between the patient's dual status and episode cost (Table 8). This association fluctuates slightly but remains

statistically significant after adding variables to account for provider-level factors, suggesting that both patient-level and provider-level factors are influential. However, there is no performance degradation observed with an increasing share of dual episodes (Table 9). Additionally, while some clinicians perform significantly worse on their dual episodes than their non-dual episodes, most clinicians perform equally well on their dual episodes as their non-dual episodes, and some even perform significantly better. This suggests that it is possible to mitigate the effect of SRFs (Table 10). Lastly, risk adjusting for dual status does not appear to substantially change the performance ranking for many providers (Table 11). Together, these results support not include a risk adjustment variable for dual enrollment in Medicare and Medicaid.

3.4.6 Method for Statistical Model or Stratification Development

To analyze the validity of the current risk adjustment model, we examined two criteria: discrimination and calibration.

1) Discrimination is a statistical criterion that evaluates the measure's ability to distinguish high-cost episodes from low-cost episodes or the ability to explain the variance in the cost of individual episodes. The amount of variance explained is estimated by the R-squared metric with the range between 0 and 1. These results are provided in Section 3.5.7.
2) Calibration evaluates the consistency of the measure in estimating episode cost across the full range of resource use patterns in the population. Calibration is estimated by the average predictive ratios across groups within the population, specifically groups partitioned by deciles of expected episode cost. A well-calibrated measure should have predictive ratios close to 1.0 across all deciles. These are discussed in Sections 3.5.8 and 3.5.9.

3.4.7 Statistical Risk Model Discrimination Statistics

The overall R-squared for the Cataract Removal with Intraocular Lens (IOL) Implantation cost measure, calculated by dividing explained sum of squares by the total sum of squares is 0.8. The adjusted R-squared is also 0.8. More information on discrimination testing for the CMS-HCC model can be found in Pope et al. 2011.[20]

3.4.8 Statistical Risk Model Calibration Statistics

The predictive ratio is calculated using the formula of average expected cost / average observed cost for all episodes in each decile.

3.4.9 Statistical Risk Model Calibration – Risk Decile

Analysis of predictive ratios by risk decile for the measure shows minimal variation among risk deciles, as predictive ratios are 1 across all risk deciles.

Table 12: Predictive Ratio by Decile of Predicted Episode Cost

Decile	Average Predictive Ratio
Decile 1	1.00
Decile 2	1.00
Decile 3	1.00
Decile 4	1.00
Decile 5	1.00
Decile 6	1.00

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

Decile	Average Predictive Ratio
Decile 7	1.00
Decile 8	1.00
Decile 9	1.00
Decile 10	1.00

3.4.10 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.[21] As noted in Section 3.5.6 and 3.5.7, these results should be interpreted alongside service assignment rules, that remove clinically unrelated services.

The remaining unexplained variance is due to variation in factors not adjusted for by the measure, such as the clinician's performance. The objective of a cost measure is to evaluate and differentiate the performance of clinicians. Therefore, achieving high explained variance is optional because the measure should only adjust for some variations in the cost of care. In collaboration with the experts from our clinical workgroup, this measure only adjusts for factors deemed outside the reasonable influence of clinicians. The service assignment rules provide context for which costs are included in the measure and which are not.

Table 12 shows that the risk adjustment model is consistent, with the average predictive ratios observed to be close to 1.00 across all deciles, maintaining a consistent ratio of 1.0. Overall, the risk adjustment model does not over- or under-predict cost across the full range of resource use patterns in the population.

3.5 Identification of Meaningful Differences in Performance

3.5.1 Method

To identify meaningful differences in performance, this analysis first examines the distribution of the measure score to highlight the performance gap between the most and least efficient clinicians. Then, this analysis examines the rate of adverse events that may occur during an episode of care to highlight the variation in frequency and cost of those events.

3.5.2 Statistical Results

Table 1 shows the distribution of the measure score at the TIN and TIN-NPI levels. There is a slight difference in the mean score for TIN and TIN-NPI levels because each level has its own attribution rules, resulting in different populations of episodes used for measure score calculation (Table 1).

While few episodes include clinically related emergency department (ED) visits (0.13%) and major procedures (0.40%), these episodes are much more costly compared to overall observed and risk-adjusted costs. For example, episodes with ED visits have a mean observed cost of $3,486 and episodes with major procedures have a mean observed cost of $4,226; the overall mean observed cost is $3,170.

3.5.3 Interpretation

There is substantial variation observed in the measure score in both TIN and TIN-NPI levels, indicated by the interquartile ranges, standard deviations, and coefficients of variation. The

[21]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.

magnitude of the observed variation is in the thousands of dollars, indicating that there are opportunities to close the gaps between the most and least efficient clinicians.

Emergency department services are relatively rare events for the cataract removal with intraocular lens (IOL) implantation procedure, but given that it is associated with high cost, reducing emergency department visits could be areas for substantial cost improvement.

3.6 Missing Data Analysis and Minimizing Bias

3.6.1 Method

Since CMS uses Medicare claims data to calculate the Cataract Removal with Intraocular Lens (IOL) Implantation measure, Acumen expects a high degree of data completeness. To further ensure that we have complete and accurate data for each patient, Acumen excludes episodes where patient date of birth information (an input to the risk adjustment model) cannot be found in the EDB, the patient does not appear in the EDB, or the patient's death date occurs before the episode trigger date.

The Cataract Removal with Intraocular Lens (IOL) Implantation measure also excludes episodes where the patient 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 patient needed to capture the clinical risk of the patient in risk adjustment. Furthermore, Parts A and B claims data may not capture all Medicare resource use if some portion of the patient's care is covered under Medicare Part C.

3.6.2 Missing Data Analysis

The table below presents the frequency of missing data across the categories of missing data that caused episodes to be excluded from the Cataract Removal with Intraocular Lens (IOL) Implantation measure. Frequency is presented in terms of the number of episodes excluded due to missing data, as well as the cost profile of episodes with missing data compared to episodes included in the measure reporting.

As a note, the episode counts below reflect exclusion from the initial population of triggered episodes. After the missing data exclusions are applied, we apply additional exclusions, as outlined in section 3.4, to this overall patient cohort to narrow the population to only applicable episodes.

Table 13: Cost Statistics for Missing Data Category

Missing Data Categories	Episodes	Observed Cost					
		Mean	Percentile				
	#		10th	25th	50th	75th	90th
No Continuous Enrollment in Medicare Parts A and B, and Any Enrollment in Part C	91,452	$2,614	$1,678	$1,871	$2,206	$3,245	$3,862
No Main Surgeon	22	$2,085	$808	$1,647	$1,852	$2,438	$3,374
Primary Payer Other than Medicare	129,779	$2,611	$1,650	$1,877	$2,216	$3,240	$3,888

3.6.3 Interpretation

The table above (Table 13) presents three distinct missing data categories, each with its associated average costs and variability. In the first category, where individuals lack continuous enrollment in Medicare Parts A and B but are enrolled in Part C, the average cost is $2,614. There is variation in costs within this group, ranging from $1,678 at the 10th percentile to $3,862

at the 90th percentile. The second category, involving patients that were not assigned a main surgeon, represents only 22 episodes. The third category comprises cases where the primary payer was other than Medicare, representing 3,374 episodes. The mean cost in this group is $2,611, with costs ranging from $1,650 at the 10th percentile to $3,888 at the 90th percentile. These categories of episodes are excluded for data completeness reasons.

4.0 Feasibility

4.1 Data Elements Generated as Byproduct of Care Processes

The data elements used in this measure are pulled from Medicare claims. They can be based on information generated, collected, and/or used by healthcare personnel during the provision of care (e.g., diagnoses), which are then translated into the appropriate coding system (e.g., ICD-10 diagnoses, MS-DRGs) for use in Medicare claims by either the original healthcare personnel or another individual.

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 CWF maintained at the CMS Baltimore Data Center. Healthcare providers submit Medicare claims to a Medicare Administrative Contractor (MAC), which is 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 such, it is not practical to wait until all claims for a given month are finalized before calculating the measure, resulting in a trade-off between efficiency (accessing the data on time) and accuracy (waiting until most claims are finalized) when determining the duration (i.e., the “claims run-out” period) after which to pull claims data. To determine the appropriate claims run-out period, Acumen has tested 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 CMS adopts this measure for use in a program, calculation and reporting would align with the program’s reporting practices.

4.3.1.2 Missing Data

This measure requires complete beneficiary information; therefore, a small number of episodes with missing data are excluded to ensure data completeness and accurate comparability across episodes. For example, episodes where the beneficiary was not enrolled in Medicare Parts A and B for the 120 days before the episode start date are excluded from this measure. Excluding these episodes 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 excluded from the measure.

4.3.1.3 Sampling

During measure testing, Acumen noted that episodes in which the beneficiary died before the episode end date exhibited different cost distributions than other episodes. As such, this measure excludes episodes to avoid negatively impacting clinician scores.

5.0 Usability and Use

5.1 Use

5.1.1 Current and Planned Use

A previous version of this measure is currently in use in MIPS. However, this measure has been revised as part of the comprehensive re-evaluation process specifically for potential use in the cost performance category of MIPS to assess clinicians reporting as individuals or groups under a contract with CMS.

For CMS to approve this measure for use in MIPS, it must be reviewed by the Pre-Rulemaking Measure Review and Measure Set Review process (PRMR-MSR; formerly referred to as the Measure Application Partnership [MAP]) and then undergo the notice-and-rulemaking process. Given these next steps, the earliest the measure could be used in MIPS is CY 2025. If in use, CMS can then determine whether to publicly report the cost measure.

5.1.2 Feedback on the Measure by Those being Measured or Others

Throughout the Cataract Removal with Intraocular Lens (IOL) Implantation measure re-evaluation, we used an iterative and extensive process to gather feedback on the measure and its results to ensure that it can be used appropriately in the MIPS program by clinicians and clinician groups who practice in this clinical area. This process also seeks to ensure that the measured entities can understand and interpret their performance results to help support decision-making. A couple of the main ways we gathered input was through reoccurring Clinician Expert Workgroup meetings, which incorporated feedback from the patient and caregiver perspective, empirical data, and discussion between clinician experts who recommend measure specifications, and through public comment periods for the measures.

5.1.2.1 Technical Assistance Provided During Development or Implementation

Clinician Expert Workgroup Meetings
For each Clinician Expert Workgroup meeting, Acumen provided empirical data (e.g., analyses on potentially relevant revisions for the measure) to inform the Clinician Expert Workgroup members' recommendations. These analyses were conducted using all administrative claims data for Medicare Parts A and B. This data was shared with Workgroup members to help inform their feedback on the measure specifications throughout its re-evaluation to ensure that the measure is appropriately assessing costs for these clinicians.

Public Comment Period
Additionally, Acumen and CMS provided two public comment periods to gather feedback the measure's re-evaluation. The first public comment period was held from February 25, 2022 to May 28, 2022, to identify which measures in use in MIPS require re-evaluation and potential revisions to those measures. A second public comment period was head in February 2023, where interested parties were invited to submit feedback via an online survey on the potential revision before consideration of their potential use in the cost performance category of the MIPS. During this feedback period, interested parties had the opportunity to view (i) measure specifications documentation, (ii) measure testing forms, (iii) clinician expert workgroup meeting summaries, and (vi) summaries of previous Wave 1 measure feedback.

5.1.2.2 Technical Assistance with Results

Clinician Expert Workgroup Meetings
Acumen provided data before or during each of the Clinician Expert Workgroup Meetings: The Comprehensive Reevaluation Webinar, and Post-Feedback Refinement Webinar. During the

meetings, Acumen would guide Workgroup members through these analyses, providing clinical and programmatic context when needed. Using this iterative process, the Workgroup members discussed the testing results in depth during each meeting and allowed the data to inform their recommendations for measure specifications. The goal was to ensure that the measure appropriately assessed clinicians' cost of care within their reasonable influence without creating potential unintended consequences so that it could be usable in the MIPS program.

Public Comment Periods

During the February 2023 public comment period, interested parties provided feedback on the appropriateness of the measures and the usability of the data. The public comments were summarized and considered the Clinician Expert Workgroup when recommending further refinements to the measures.

Education and Outreach

Acumen directly conducted outreach via email to tens of thousands of interested parties using a contact list developed through previous public engagement efforts, as well as CMS and Quality Payment Program (QPP) listservs. Acumen also contacted specialty societies that may have interest in these measures due to the types of clinicians that they represent.

Acumen worked closely with QPP Service Center to respond to stakeholder inquiries during the public comment period and continued to answer questions after the period ended.

5.1.2.3 Feedback on Measure Performance and Implementation

Clinician Expert Workgroup Meetings

Feedback from the Workgroup members were recorded throughout the meeting. More formal feedback was gathered using polls, typically requesting for votes on certain specifications or appropriateness of the measure. These polls were conducted following each meeting and on an ad hoc basis, as needed.

Public Comment Periods

For the 2022 public comment period, Acumen received 20 comments and for the 2023 public comment period, Acumen received 18 comments. These responses included comments from specialty societies representing large numbers of potentially attributed clinicians and from individuals.

Survey responses were collected via an online survey, which contained general and detailed questions on the measure specifications.

5.1.2.4 Feedback from Measured Entities and Other Entities

Public Comment Periods

The MACRA Episode-Based Cost Measures: Comprehensive Reevaluation Public Comment Summary Report presents interested parties' feedback from the initial public comment period in 2022.[22] The 2023 Revised Cost Measure Feedback Period Summary Report presents stakeholder feedback gathered during the second public comment period.[23] The measure-specific feedback was used as the basis for refinements that were made to the measures. See Section 5.1.2.5 for refinements made to the Cataract Removal with Intraocular Lens (IOL) Implantation measure.

---

[22] CMS, "MACRA Episode-Based Cost Measures: Comprehensive Reevaluation Public Comment Summary Report," Cost Measures Feedback Page, https://www.cms.gov/files/document/wave-one-public-comment-summary-report.pdf.

[23] CMS, "2023 Revised Cost Measure Feedback Period Summary Report," Cost Measures Feedback Page, https://www.cms.gov/files/document/2023-revised-cost-measure-feedback-period-summary-report.pdf.

5.1.2.5 Consideration of Feedback

Public Comments

Careful consideration was given to all feedback gathered through public comment, and several updates were made to the measure based on the recommendations of commenters and the Clinician Expert Workgroup comprised of subject matter and measure-development experts. Acumen conducted analyses into potential adjustments that could be made to the measures to improve their ability to assess the intended clinician population.

After public comment periods, Acumen compiled the feedback and provided the Clinician Expert Workgroup this information, along with the empirical analyses, to inform recommendations for any refinements needed to ensure that the measure is capturing what it was intended to capture.

The changes to the Cataract Removal with Intraocular Lens (IOL) Implantation measure made through re-evaluation include:

- Expand the measure scope to include patients with certain ocular conditions, such as macular degeneration, glaucoma, and diabetic eye disease
- Include the costs of additional clinically related services, such as pre-operative testing, additional telehealth services, durable medical equipment (DME), emergency department (ED) visits for ocular complaints, and durable medical equipment

5.2 Usability

5.2.1 Improvement

The version of the measure has not yet been implemented, and as such has not had influence over performance. Our testing suggests that there is a sufficiently large difference in measure scores among clinicians to meaningfully determine a difference in performance. The potential for this measure to distinguish between good and poor performance is promising in its ability to encourage improvement in cost efficient care.

5.2.2 Unexpected Findings

There were no unexpected findings during the development and testing of this measure. This version of the measure has not been implemented at this time, so we do not have data that confirms unexpected findings related to its implementation.

However, Acumen did consider potential unintended consequences of having a cost measure for this clinical area (e.g., potential stinting in care to receive a better cost score). For example, the empiric validity data previously presented in section 3.3 demonstrates that many of the included services are not associated with the costs of adverse events, suggesting that cost improvement can be achieved without increasing occurrence of adverse events.

Additionally, CMS monitors measures that are in use and has multiple processes in place to allow for changes to a measure if appropriate. These include i) annual maintenance for non-substantial changes and upkeep, ii) ad hoc maintenance if a specific issue occurs or a large change in clinical guidance takes place, and iii) measure reevaluation every three years where the suitability of a measure's specifications is comprehensively reassessed. If in the event the measure did have any unexpected findings, it would be identified and resolved through one of these methods.

5.2.3 Unexpected Benefits

Since this version of the measure has not been implemented at this time, there are no testing results that identify unexpected benefits. However, many clinicians can only be assessed by the

MSPB Clinician and TPCC measures in the cost performance category currently. This measure would provide a more tailored assessment of the care they have influence over, which many clinicians may prefer to be measured by compared to the population-based cost measures like MSPB Clinician or TPCC.

6.0 Related and Competing Measures

6.1 Relation to Other Measures

There are no competing measures with this measure. However, the following measures have been identified as potentially related.

Table 14. Quality Measures Potentially Relevant for the Cataract Removal with Intraocular Lens (IOL) Implantation Episode Group

Measure Title	Measure ID	Measure Description	Measure Type
Cataracts: 20/40 or Better Visual Acuity within 90 Days Following Cataract Surgery	00114	Percentage of cataract surgeries for patients aged 18 and older with a diagnosis of uncomplicated cataract and no significant ocular conditions impacting the visual outcome of surgery and had best-corrected visual acuity of 20/40 or better (distance or near) achieved in the operative eye within 90 days following the cataract surgery.	Outcome
Cataracts: Patient Satisfaction within 90 Days Following Cataract Surgery	00117	Percentage of patients aged 18 years and older who had cataract surgery and were satisfied with their care within 90 days following the cataract surgery, based on completion of the Consumer Assessment of Healthcare Providers and Systems Surgical Care Survey.	Outcome
Cataract Surgery: Difference Between Planned and Final Refraction	00113	Percentage of patients aged 18 years and older who had cataract surgery performed and who achieved a final refraction within +/- 1.0 diopters of their planned (target) refraction.	Outcome

The MIPS quality measures listed above are related to the Cataract Removal with Intraocular Lens (IOL) Implantation measure as they include metrics that focus on similar patient cohorts, are clinically related to the care provided for the episode group, or assess complementary care that may not be directly captured by the cost measure.

6.2 Harmonization

During the measure's development, the Clinician Expert Workgroup specifically considered how to align relevant cost and quality measures (e.g., episode window length). This cost measure aligns with the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, and the domain of Efficiency and Cost Reduction. Through this measure, we aim to improve care by optimizing health outcomes and resource use associated with this procedure. The development of episode groups for resource use analysis is also required by section 101(f) of MACRA.

6.3 Competing Measures

There are no measures that conceptually address both the same measure focus and the same target population as the Cataract Removal with Intraocular Lens (IOL) Implantation measure. This is a revised version of the currently used Routine Cataract Removal with Intraocular Lens (IOL) Implantation, and would replace the current measure if approved for use in MIPS in the future.

Additional Information

The Cataract Removal with IOL Implantation Clinician Expert Workgroup Members:
As noted above, the following members provided detailed feedback on the measure specifications throughout its development based on public comments, clinical expertise, and empirical analyses.

David Glasser, MD, American Academy of Ophthalmology John Hitchens, CRNA, ARNP, FAANA, Hitchens and Henke
Keith Walter, MD, Wake Forest University School of Medicine/Wake Forest University Eye Center
Mitchell Jackson, MD, Jacksoneye, Lake Villa
Parag Parekh, MD, ClearView Eye Consultants Scott Friedman, MD, Florida Retina Consultants
Richard Tipperman, MD, Chair ASCRS Cataract Committee

Measure Developer Updates and Ongoing Maintenance
The measure is not currently in use, but the earliest possible release of the measure in MIPS would be CY2025. If the measure becomes finalized for use in MIPS, it would undergo annual maintenance and a comprehensive re-evaluation every 3 years. This measure is included on the 2023 Measures Under Consideration (MUC) List and will be reviewed by PRMR in winter of 2023-2024. There are no further updates or reviews for this measure scheduled at this time.