Quality Payment
PROGRAM

Melanoma Resection

Measure Testing Form

Summer 2020 Field Testing

Table of Contents

1.0 Introduction ... 3
1.1 Field Testing ... 3
1.1.1 Overview ... 3
1.1.2 Providing Feedback ... 3
2.0 Measure Testing: Importance ... 4
2.1 Evidence to Support the Measure Focus ... 4
2.1.1 Measure Description ... 4
2.1.2 Evidence for Measure Focus ... 4
2.2 Performance Gap ... 6
2.2.1 Rationale ... 6
2.2.2 Performance Scores ... 6
3.0 Scientific Acceptability ... 7
3.1 Data Sample Description ... 7
3.1.1 Type of Data Used for Testing ... 7
3.1.2 Specific Dataset Used for Testing ... 7
3.1.3 Dates of the Data Used in Testing ... 7
3.1.4 Levels of Analysis Tested ... 7
3.1.5 Entities Included in the Testing and Analysis ... 7
3.1.6 Patient Cohort Included in the Testing and Analysis ... 7
3.1.7 Social Risk Factors Included in Analysis ... 9
3.2 Validity Testing ... 9
3.2.1 Level of Validity Testing ... 9
3.2.2 Method of Validity Testing ... 9
3.2.3 Statistical Results from Validity Testing ... 11
3.2.4 Interpretation ... 13
3.3 Exclusions Analysis ... 13
3.3.1 Method of Testing Exclusions ... 13
3.3.2 Statistical Results from Testing Exclusions ... 14
3.3.3 Interpretation ... 14
3.4 Risk Adjustment or Stratification ... 15
3.4.1 Method of Controlling for Differences ... 15
3.4.2 Conceptual, Clinical, and Statistical Methods ... 17
3.4.3 Conceptual Model of Impact of Social Risks ... 18
3.4.4 Statistical Results ... 18
3.4.5 Analyses and Interpretation in Selection of Social Risk Factors ... 18
3.4.6 Method for Statistical Model or Stratification Development ... 20
3.4.7 Statistical Risk Model Discrimination Statistics ... 20
3.4.8 Statistical Risk Model Calibration Statistics ... 20
3.4.9 Statistical Risk Model Calibration – Risk Decile ... 21
3.4.10 Interpretation ... 21
3.5 Identification of Meaningful Differences in Performance ... 22
3.5.1 Method ... 22
3.5.2 Statistical Results ... 22
3.5.3 Interpretation ... 25
3.6 Missing Data Analysis and Minimizing Bias ... 26
3.6.1 Method ... 26
3.6.2 Missing Data Analysis ... 26
3.6.3 Interpretation ... 27
Contact Information ... 28

1.0 Introduction

This Measure Testing Form provides results for the testing of the Melanoma Resection measure that is being field tested between August 17 and September 18, 2020. Section 2 describes the scientific literature to support the measure as well as evidence of a performance gap among clinicians and clinician groups. Section 3 presents testing information and results for the measure.

The testing form accompanies the draft Measure Methodology document and draft Measure Codes List file posted on the MACRA Feedback Page,[1] which comprise the draft specifications for the Melanoma Resection measure.

1.1 Field Testing

1.1.1 Overview

As a part of the measure development process, field testing is an opportunity for clinicians and other stakeholders to learn about episode-based cost measures and provide input on the draft measure specifications. During field testing, Field Test Reports are distributed on the Quality Payment Program website[2] for group practices (identified by Tax Identification Number [TIN]) and individual clinicians (identified by combination of TIN and National Provider Identifier [NPI]) who meet the minimum number of cases for each measure. A volume threshold of 10 episodes was used for procedural and acute inpatient medical condition episode groups (including Melanoma Resection), and 20 episodes for chronic condition episode groups. Draft measure specifications and supplemental documentation are available on the MACRA Feedback page.[3] Stakeholder feedback during field testing is collected on the draft specifications for each measure.

1.1.2 Providing Feedback

The feedback from field testing helps inform refinements to the measures before Centers for Medicare and Medicaid Services (CMS) considers them for potential use in the Cost performance category of the Merit-based Incentive Payment System (MIPS). Acumen is collecting stakeholder feedback on the draft measure specifications of the 5 episode-based cost measures during the field testing period, between August 17 and September 18, 2020, through this online Field Testing Feedback Survey.[4]

Specific questions about the draft Melanoma Resection measure specifications are available in the Questions for Field Testing Measure Specifications document,[5] which stakeholders can use as a reference while reviewing the field testing materials.

[1] CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback
[2] CMS, "QPP Account," Quality Payment Program, https://qpp.cms.gov/login.
[3] CMS, "Cost Measure Field Testing", MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-MIPS-and-APMs.html.
[4] The field testing online survey will be open beginning August 17, 2020 at this link: https://www.surveymonkey.com/r/2020-cost-measures-field-testing.
[5] This document will be available on the MACRA Feedback Page once field testing begins. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-MIPS-and-APMs.html

2.0 Measure Testing: Importance

2.1 Evidence to Support the Measure Focus

2.1.1 Measure Description

The Melanoma Resection cost measure evaluates clinicians and clinicians groups' risk-adjusted cost to Medicare for patients who undergo an excision procedure to remove a cutaneous melanoma. The measure score is a clinician or clinician group's average risk-adjusted cost across all attributed episodes for the episode group. This procedural measure includes services that are clinically related and under the reasonable influence of the attributed clinician or clinician group during the 30 days prior to the melanoma resection procedure which opens or "triggers" the episode and in the 90 days after the procedure. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure.

2.1.2 Evidence for Measure Focus

A recent study indicates that clinician beliefs about treatment and the efficacy of particular therapies may be the most important factors explaining the variation in health care expenditures.[6] However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to inform clinicians on the costs associated with their decision-making and to incentivize cost-effective, high-quality care. A cost measure offers 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 care quality through changes in clinical practice.

According to the literature and feedback received through stakeholder input activities, this measure's focus represents an area where there are opportunities for improvement. As discussed in the rest of this section, primary opportunities for improving melanoma resection cost outcomes include selectively performing sentinel lymph node (SLN) biopsies, performing follow-up procedures as close as possible together, and reducing downstream complications through adherence to clinical guidelines.

More selectively performing SLN biopsies (i.e., not performing them for all melanomas) will allow for cost savings due to fewer procedural costs for SLN services, as well as fewer complications due to additional procedures. The SLN biopsy is a procedure frequently performed after a melanoma excision, when the size of the melanoma indicates potential disease spread. The SLN biopsy can confirm presence and absence of metastases, where a positive SLN biopsy result indicates non-localized and thus more severe disease (regional/distant). Performing this procedure may assist in long-term disease-free survival, as the SLN procedure can identify nodal metastases that would be otherwise caught later via other methods of observation, such as monitoring. This is corroborated by one study that found that the 10-year disease-free survival rate for those that underwent an SLN biopsy were significantly greater than those that were simply monitored for melanoma recurrence in lymph nodes.[7] However, due to costs and the nearly triple complication rate of SLN biopsies compared to

[6] 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.

[7] Sondak, V. K., MD. (2014, April 22). Long-Term Outcomes Support Sentinel-Node Biopsy for Staging Melanoma. Retrieved July 23, 2020, from https://www.onclive.com/view/long-term-outcomes-support-sentinel-node-biopsy-for-staging-melanoma

standard melanoma resections, these procedures should only be used selectively.[8,9] One meta-analysis suggested that early-stage melanomas (Stage 1 or less) should not be followed routinely by an SLN biopsy, as the risks outweigh the benefits of the procedure. At an early stage, the patient is less likely to have metastases to be removed, meaning that the only benefits are confirming no disease presence, while adding risk through exposing the patient to potential surgical complications.

As another example of opportunity for improvement, research suggests that the timing of primary excision and secondary reconstructive procedures may contribute to better cost outcomes. One study indicated that roughly 20% of Medicare patients undergoing surgical treatment to remove a melanoma experience a delay of longer than 1.5 months between biopsy and excision procedure, with longer delays being correlated with higher morbidity and mortality compared to excising within 30 days of biopsy.[10] This is especially true for early-stage melanomas, which are the intended scope of the Melanoma Resection measure. Minimizing delays between the excision of the melanoma and reconstructive procedures may also lead to opportunities for improvement. One study suggested that performing the reconstructive procedure immediately after excision can generate substantial savings, especially in the inpatient setting, noting a 38.5% lower cost in the treatment arc with immediate reconstruction compared to the cost of delayed reconstruction. While immediate reconstruction could potentially allow a reconstructed wound to have residual disease, the paper notes an acceptably low rate of residual tumors requiring operation.[11]

Beyond timing of procedures, focusing on other ways to reduce downstream complications relevant to the index melanoma resection presents an opportunity to lower the cost of care. These complications can include surgical site infections (SSIs), delayed wound healing or wound dehiscence, as well as skin grafts or skin substitutes. While clinical characteristics may predispose certain patients to SSIs, the likelihood of an SSI can be reduced through evidence-based practices. These practices include, but are not limited to, proper administration of any necessary antibiotics and appropriate use of medical and sanitary equipment by the medical staff, including wearing proper surgical attire and disinfecting the surgical site prior to excision.[12,13,14] Adhering to these evidence-based practices mitigates common post-operative

---

[8] Arguello-Guerra, Lilia, Estefanía Vargas-Chandomid, Jose Manuel Díaz-González, Silvia Méndez-Flores, Ana Ruelas-Villavicencio, and Judith Domínguez-Cherit. “Incidence of Complications in Dermatological Surgery of Melanoma and Non-Melanoma Skin Cancer in Patients with Multiple Comorbidity and/or Antiplatelet-Anticoagulants. Five Year Experience in Our Hospital.” Cirugía y Cirujanos (English Edition) 86, no. 1 (May 2019). https://doi.org/10.24875/cirue.m18000003.

[9] Moody, J., R. Ali, and J. Hardwicke. “Complications of Sentinel Lymph Node Biopsy for Melanoma - A Systematic Review of the Literature.” International Journal of Surgery 36 (June 22, 2016). https://doi.org/10.1016/j.ijsu.2016.08.326.

[10] Lott JP, Narayan D, Soulos PR, Aminawung J, Gross CP. Delay of Surgery for Melanoma Among Medicare Beneficiaries. JAMA Dermatol. 2015;151(7):731–741. https://doi.org/10.1001/jamadermatol.2015.119.

[11] "Oncology; Division of Surgery and Oncology Reports Findings in Melanomas (Melanoma Extirpation with Immediate Reconstruction: The Oncologic Safety and Cost Savings of Single-Stage Treatment)." 2016. Medical Devices & Surgical Technology Week, Aug 14, 159. https://search.proquest.com/docview/1808815332?accountid=165523.

[12] “Surgical Site Infections.” Surgical Site Infections | Johns Hopkins Medicine. Johns Hopkins. Accessed May 1, 2020. https://www.hopkinsmedicine.org/health/conditions-and-diseases/surgical-site-infections.

[13] “Surgical Site Infection (SSI) Prevention .” The Johns Hopkins Hospital. Johns Hopkins Medicine, July 1, 2012. https://www.hopkinsmedicine.org/heic/docs/SSI_prevention_best_practices_summary.pdf.

[14] Berríos-Torres SI, Umscheid CA, Bratzler DW, et al. Centers for Disease Control and Prevention Guideline for the Prevention of Surgical Site Infection, 2017. JAMA Surg. 2017;152(8):784–791. https://doi.org/10.1001/jamasurg.2017.0904.

complications, such as infection and wound dehiscence, which require follow-up and additional treatment outside of the standard treatment arc of a melanoma resection.

2.2 Performance Gap

2.2.1 Rationale

Given the incidence and mortality of melanoma in the Medicare-aged population, the Melanoma Resection measure represents an opportunity to control Medicare spending for melanoma as the incidence of melanoma rises. In the United States, the average age when melanoma is diagnosed is 65, with incidence and melanoma-specific mortality increasing with age and peaking in those aged 65-74 years.[15] It is estimated that 196,060 cases of melanoma will be newly diagnosed in 2020. Furthermore, it is estimated that the total annual treatment costs for melanoma are $3.3 billion annually, a figure that is anticipated to continue to rise due to the increasing incidence of melanoma.[16]

The Melanoma Resection episode-based cost measure was recommended for development by an expert clinician committee—the Dermatologic Disease Management Clinical Subcommittee. Based on the initial recommendations from the Clinical Subcommittee, the subsequent measure-specific Clinician Expert Workgroup provided extensive, detailed input on this measure.

2.2.2 Performance Scores

To demonstrate the performance gap captured in the measure, Table 1 below presents a distribution of performance scores for 1,812 clinician group practices and 2,188 practitioners attributed episodes in 2019. These counts represent attributed clinicians and clinician groups billing Part B Physician/Supplier claims under a 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.

Table 1. Distribution of Performance Scores

Metric	TIN	TIN-NPI
Mean score	$1,520	$1,539
Score Interquartile Range (IQR)	$443	$482
Score Percentile		
10th	$1,082	$1,067
25th	$1,251	$1,244
50th	$1,436	$1,451
75th	$1,694	$1,726
90th	$2,049	$2,132

[15] Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2016, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2016/, based on November 2018 SEER data submission, posted to the SEER web site, April 2019.

[16] "Skin Cancer Facts & Statistics: What You Need to Know "Skin Cancer Facts and Statistics. Skin Cancer Foundation. Accessed May 1, 2020. https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

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), Medicare Enrollment Database (EDB), Common Medicare Environment (CME), and United States Census Bureau's American Community Survey (ACS).

3.1.2 Specific Dataset Used for Testing

The Melanoma Resection 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 adjustors. Episode costs are payment standardized and risk adjusted to ensure accurate comparison of cost across clinicians. Payment standardization adjusts the allowed amount for a Medicare service to limit observed differences in costs to those that may result from health care delivery choices. Data from the EDB are used 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.

For measure testing, data from the ACS and CME are used in analyses evaluating social risk factors in risk adjustment.

3.1.3 Dates of the Data Used in Testing

Melanoma Resection episodes ending from January 1, 2019, through December 31, 2019.

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

The overall population used for testing includes 18,464 clinician group practices and 47,139 practitioners, which includes any clinician groups/practitioners who had at least one Melanoma Resection episode in the measurement period. After applying exclusions and the case minimum, the final population for testing and analyses included 1,812 clinician group practices and 2,188 practitioners who were attributed 10 or more Melanoma Resection episodes during the measurement period. Episodes from all 50 states and the District of Columbia triggered in the following settings were included:

- Ambulatory surgical centers (ASC)
- Hospital outpatient department (HOPD)
- Ambulatory/office-based care

3.1.6 Patient Cohort Included in the Testing and Analysis

65,980 Medicare patients, with a mean age of 75.96, (from 77,945 episodes) were included in measure testing and analyses (where patient populations are not subject to any case volume restrictions).

The patient population for the Melanoma Resection measure calculation consists of Medicare beneficiaries enrolled in Medicare Parts A and B (but not Part C) who undergo an excision

procedure to remove a cutaneous melanoma that triggers a Melanoma Resection episode, as identified by trigger Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) codes on Part B claims for a either a cutaneous excision or tissue rearrangement. This CPT/HCPCS code must be accompanied by an International Classification of Diseases, 10th Edition (ICD-10) diagnosis code for melanoma (C43 or D03) on the trigger claim in order to trigger an episode.

The exclusion criteria are:

- The patient did not have 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 patient was not 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 patient does not have a sufficient 120-day lookback period.
- The patient date of birth is missing.
- The patient death date occurred before episode end.
- The episode has no attributed clinician or clinician group.
- The episode trigger claim was not in an outpatient (OP), inpatient (IP), or ASC setting.
- The episode contains a Mohs surgery procedure code accompanied by a melanoma diagnosis code.
- The episode trigger claim does not have a melanoma diagnosis code (C43 or D03).

To determine whether the Melanoma Resection measure's exclusion 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 exclusion criteria, (ii) episodes without exclusion criteria, (iii) patients with exclusion criteria, and (iv) patients without exclusion criteria.

This analysis shows that the Melanoma Resection measure's exclusion criteria have a minimal effect on the percentage of patients in any particular demographic category. Results show one exception for the prevalence of HCC12 for "Breast, Prostate, and Other Cancers and Tumors" which is 59.4 percentage points greater in the measure population after exclusion criteria is applied. This is because the Melanoma Resection exclusion criteria specifically excludes patients who undergo an excision without a melanoma diagnosis, resulting in a much higher share of patients with a cancer diagnosis (captured under HCC12) after exclusions, compared to the baseline population of overall excision procedures.

Across all other dimensions, the difference between patients included and excluded in the measure is less than 4.9 percentage points across each of the characteristics in the analysis at TIN level testing, and less than 5.2 percentage points at TIN-NPI level testing. The largest difference observed is in the breakdown of male and female patients which shifts slightly after applying the exclusion criteria: the breakdown is 41.4% female and 58.6% male without applying the exclusion criteria, compared to 36.5% female and 63.5% male with exclusion criteria at the TIN level, and 36.2% female and 63.8% male at the TIN-NPI level. This general breakdown between male and female aligns with findings from the American Academy of Dermatology that indicate that men are twice as likely at age 65 and 3 times as likely at age 80 to develop melanoma compared to women.[17] Shifts in other demographic categories are less pronounced. To illustrate, 18.6% of patients are between the ages of 65 to 69 before applying the exclusion criteria, compared to 22.0% after applying exclusion criteria at TIN level and 21.5% after applying exclusion criteria TIN-NPI level. Additionally, the differences in the

[17] "Melanoma Strikes Men Harder." American Academy of Dermatology. Accessed July 8, 2020. https://www.aad.org/public/diseases/skin-cancer/types/common/melanoma/men-50.

percentage of patients in each race category are all less than 0.6 percentage points with and without the exclusion criteria. Overall, these results indicate that there is minimal shift in patient characteristics as a result of using the exclusion criteria listed above at both TIN and TIN-NPI level testing.

3.1.7 Social Risk Factors Included in Analysis

The social risk factors analyzed were variables from the ACS, EDB, and CME. ACS variables are either at the Census Block Group or Zone Improvement Plan (ZIP) Code level. Social risk variables analyzed include the following:

- Race (EDB)
  - Asian, Black, Hispanic, North American Native, White, and Other
- Sex (EDB)
  - Female, male
- Dual status (CME)
  - Full dual, partial dual, non-dual to indicate whether a patient is dually enrolled in Medicare and Medicaid
- 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%
- Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) Index (ACS)
  - Continuous variable (composite score of multiple community-level metrics, such as property values, density of living spaces, and poverty level) that can theoretically range from 0 to 100[18]

3.2 Validity Testing

3.2.1 Level of Validity Testing

Our performance measure score validity testing included systematic assessment of both face validity and empirical validity testing.

3.2.2 Method of Validity Testing

Face Validity

The Melanoma Resection measure was developed through a structured, iterative process for gathering detailed input from recognized clinician experts on the measure. 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

[18] Refer to Section 3, page 42 of this AHRQ publication for the scoring algorithm used to calculate the AHRQ SES index variable

attributed clinician for a defined patient population (i.e., the ability of the measure score to differentiate good from poor performance).

In developing this measure, Acumen incorporated input from:

(i) a Dermatologic Disease Management Clinical Subcommittee;
(ii) a Melanoma Resection Clinician Expert Workgroup;
(iii) a Technical Expert Panel (TEP); and
(iv) the Person and Family Committee (PFC).

This process is detailed in the Episode-Based Cost Measures Development Process document posted on the MACRA Feedback Page.[19]

One of the key roles of the measure-specific Clinician Expert 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 resecting a cutaneous melanoma, thus limiting cost variation unrelated to clinician care this measure. Assigned services occurring in the outpatient, clinician, and ambulatory service setting were defined separately for the pre- and post-trigger periods and include only services directly related to evaluation, testing, treatment, or follow-up for an excision procedure to remove a cutaneous melanoma and relevant complications. Home health, emergency department, and inpatient services are assigned only in the post-trigger period so as to capture downstream services that the patient might require related to the excision procedure.

Empirical Validity Testing

We undertook 2 approaches to estimate the measure's validity. In the first approach, we evaluated the empirical validity of the Melanoma Resection measure by examining correlation with known indicators of resource or service utilization based on a literature review, specifically complications related to resecting a cutaneous melanoma. For this analysis, we compared the ratio of observed to expected spending at the provider level for Melanoma Resection episodes with and without complications occurring in the post-trigger period. This analysis sought to confirm the expectation that the Melanoma Resection measure captures variation in service utilization as an indicator of clinician cost performance. We expect episodes with complications related to the trigger procedure would have higher observed to expected (O/E) cost ratios, since these services should yield higher cost, even after accounting for patient clinical characteristics via risk adjustment. Conversely, episodes without these downstream costs should have lower O/E cost ratios, demonstrating that the measure can differentiate good from poor cost performance.

In the second approach, we evaluated how different types of cost impact measure scores. To define types of cost, services or costs included in the Melanoma Resection measure were classified into clinically coherent groups of services, called "clinical themes." The Melanoma Resection measure clinical themes are:

- **Primary Resection**: Includes initial trigger procedure services and any associated inpatient hospitalizations.
- **Secondary Excision**: Includes services related to secondary excision after the initial trigger resection procedure, including associated office and emergency department visits and inpatient hospitalizations.

[19] CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

- **Secondary Reconstruction**: Includes services related to secondary reconstruction after the initial trigger resection procedure, including associated office and emergency department visits and inpatient hospitalizations.
- **Lymph Node Services**: Includes lymph node biopsy, excision, and intraoperative lymph node imaging services.
- **Infection**: Includes services related to post-surgical infection, including associated office and emergency department visits, inpatient hospitalizations, post-acute care, and associated imaging or laboratory testing.
- **Wound Care**: Includes services related to wound healing (including disruption and impairment) and wound care, including skin substitutes and wound repair aids as well as associated office and emergency department visits, inpatient hospitalizations, and home health services.
- **Other Surgical Complications**: Includes inpatient, outpatient, emergency department, and home health care for other unspecified surgical complications.
- **Other Inpatient hospitalization**: Includes other melanoma-related inpatient, outpatient, and inpatient rehabilitation facility services not covered in another clinical theme.
- **Other Imaging**: Includes other melanoma-related imaging tests not covered in another clinical theme.
- **Other ED visits**: Includes other melanoma-related emergency department services, including physician Part B services for ED care.
- **Other Post-Acute Care**: Includes melanoma-related physician-billed skilled nursing facility care, including post-acute care (PAC).
- **Other Home Health Services**: Includes melanoma-related home health care services not covered in another clinical theme, including home health and assisted living facility care and associated physician visits.
- **Other Pre-Operative Outpatient Services**: Includes other melanoma-related outpatient services before the trigger resection procedure, not covered in another clinical theme.
- **Other Post-Operative Outpatient Services**: Includes other melanoma-related outpatient services after the trigger resection procedure, not covered in another clinical theme.

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 calculated the Pearson correlation between the cost of each clinical theme and the overall risk-adjusted cost for an episode.

We expect that clinical themes related to post-trigger complications, such as Infection, Secondary Reconstruction, and Other Post-Operative Services, should have the highest correlation with risk-adjusted cost, as complications are likely associated with high costs even after accounting for patient characteristics. We would expect to see similar positive correlations with risk-adjusted cost for other clinical themes representing further downstream complications, such as Secondary Excision and Wound Care. Conversely, we would expect weaker correlations with risk-adjusted cost for clinical themes relevant to preventative and diagnostic services, such as Other Pre-Operative Outpatient Services and Lymph Node Services.

3.2.3 Statistical Results from Validity Testing

Table 2 below presents the results from the first analysis of validity. The mean O/E ratio for all episodes is 1.02. The mean O/E ratio for episodes with services related to complications during the post-trigger period is 1.86, compared with 0.80 for episodes without services relating to complications during the post-trigger period. Additionally, there is greater variation in O/E ratio among episodes with complications than episodes without complications.

Table 2: Distribution of Observed to Expected Ratios

Episode Type	Observed / Expected Ratio										
	Mean	Std. Dev.	Percentile								
			1st	5th	10th	25th	50th	75th	90th	95th	99th
All Final Episodes	1.02	0.86	0.19	0.33	0.46	0.65	0.83	1.10	1.63	2.24	5.01
Episodes with Services Related to Melanoma Resection Complications	1.86	1.44	0.47	0.67	0.79	1.06	1.43	2.05	3.34	4.79	8.41
Episodes without Services Related to Melanoma Resection Complications	0.80	0.38	0.18	0.30	0.41	0.61	0.76	0.94	1.14	1.31	1.95

Table 3 below presents a subset of the results from the clinical themes analysis. The results demonstrate the greatest correlation between clinical themes associated with post-operative complications, such as Infection (correlation: 0.72) and Secondary Reconstruction (correlation: 0.46) and risk-adjusted cost. Other clinical themes related to post-operative complications and procedures followed with modest correlations with risk adjusted cost, such as the Other Post-Operative Outpatient Services (correlation: 0.35) and the Surgical Complications (correlation: 0.31) clinical themes. The Secondary Excision (correlation: 0.22) clinical theme falls in between themes capturing complications and those related to diagnostic procedures and preparatory/preventative services. The Secondary Excision clinical theme contains services related to the trigger procedure, such as performing an additional excision to attain local disease control, but also services like post-operative office visits, which are more diagnostic in nature. Clinical themes primarily built around diagnostic/preventative services, such as the Other Pre-Operative Outpatient Services (correlation: 0.07) and Lymph Node Services (correlation: 0.20) themes, had lower correlation with risk-adjusted cost.

Table 3: Clinical Themes

Clinical Theme	Pearson Correlation With Risk-Adjusted Cost
Secondary Excision	0.22
Secondary Reconstruction	0.46
Lymph Node Services	0.20
Infection	0.72
Wound Care	0.22
Other Surgical Complications	0.31
Other Emergency Department Visits	0.26
Other Pre-Operative Outpatient Services	0.07
Other Post-Operative Outpatient Services	0.35

3.2.4 Interpretation

As expected, the average O/E cost ratio for episodes with post-trigger complications is higher than for episodes without downstream complications. This result demonstrates that the Melanoma Resection measure is able to accurately capture higher resource use.

The clinical themes analysis demonstrates that high risk adjusted cost is more strongly associated with themes related to complications and post-trigger follow-up procedures, such as the Infection and Secondary Reconstruction clinical themes. Furthermore, clinical themes associated with pre-operative services and other diagnostic services, such as Other Pre-Operative Outpatient Services and Lymph Node Services, are not as strongly correlated with risk-adjusted cost. This indicates that the measure may disincentivize higher rates of post-operative complications, such as secondary reconstructions, infection, and other surgical complications, without disincentivizing the administration of pre-operative and diagnostic care, such as imaging, laboratory tests, and physician visits, where appropriate.

Additionally, we see evidence that theme correlation with cost does not come only from a mechanical increase in episode costs from high-cost service categories. For example, correlation with risk-adjusted cost is strong not only for higher cost themes, such as Secondary Reconstruction (correlation: 0.46, average cost: $1,236), but also for lower cost themes such as Infection (correlation: 0.72, average cost: $596). Furthermore, the Infection theme is substantially more strongly correlated with risk-adjusted cost than the highest-cost clinical theme, Lymph Node Services (correlation: 0.20, average cost $3,281). This provides further credence that clinical themes associated with complications accurately capture clinical cost variation agnostic of their average cost, while not disincentivizing providers from providing diagnostic services.

3.3 Exclusions Analysis

3.3.1 Method of Testing Exclusions

Exclusions are used in the Melanoma Resection measure to ensure a comparable patient population within the scope of the measure's focus on the surgical removal of cutaneous melanomas 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 exclusions added to ensure a homogenous patient population.

- Episodes where patient death date occurred before the episode end date
    - These episodes were excluded as they may not accurately reflect a clinician's performance. Episodes where the patient 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 without a melanoma diagnosis accompanying the trigger code on the trigger claim
    - Episodes were excluded if patients underwent a cutaneous excision or tissue rearrangement unrelated to a melanoma, as these episodes would be outside the intended scope of this measure.
- Episodes containing a Mohs surgery procedure accompanied by a melanoma diagnosis code
    - These episodes were excluded as expert clinical input indicated that Mohs surgery is not typically used to treat cutaneous melanoma solvable by resection.

As Mohs surgery is not an excision procedure, these episodes fall outside of the intended scope of the measure to capture services related to resecting a melanoma.

- Episodes where the trigger claim is for a procedure that was not performed in an office, OP, IP, or ASC setting
    - Expert clinical input suggested that episodes performed outside these settings would have different costs than the intended measure population of interest.

Given the rationales for these exclusions, we would 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 and ratio of observed over expected spending (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 2 patient cohorts. A full list of the exclusions used for the Melanoma Resection measure is provided in the draft Measure Codes List available on the MACRA Feedback Page.[20]

3.3.2 Statistical Results from Testing Exclusions

Table 4 below presents observed cost statistics and O/E cost ratios for the Melanoma Resection measure exclusions. Cost statistics are also provided for the set of final episodes included in the Melanoma Resection measure for comparison, with a testing volume threshold of 10 episodes at the TIN and TIN-NPI levels. For the standard exclusion 'Not in OP, IP, or ASC setting,' in the table below, this patient cohort is excluded from the measure in order to assess episodes in the intended setting.

Table 4: Cost Statistics for Measure Exclusions

Exclusion	Episodes		Observed Cost			O/E Cost Ratio		
			Mean	Percentile		Mean	Percentile	
	#	%		10th	90th		10th	90th
All Episodes Meeting Triggering Logic	**83,341**	**100.00%**	**$1,925**	**$432**	**$4,233**	**1.19**	**0.46**	**1.76**
Patient Death in Episode	460	0.55%	$3,404	$450	$6,788	1.83	0.41	2.98
Not in OP, IP, or ASC Setting	12	0.01%	$2,239	$343	$6,292	0.85	0.30	1.26
Mohs Surgery	3,343	4.01%	$1,717	$322	$3,310	1.09	0.32	1.77
Final Episodes (TIN)	64,728	77.67%*	$1,895	$434	$4,237	1.17	0.48	1.73
Final Episodes (TIN-NPI)	46,418	55.70%*	$1,867	$449	$4,199	1.16	0.49	1.75

*not all exclusions are listed in this table

3.3.3 Interpretation

The statistical results show that excluded episodes differ in average O/E cost ratio and the cost ratio distribution when compared to the final episode populations, supporting the exclusion of these episodes to ensure a comparable patient cohort that will yield meaningful information to attributed clinicians. Further discussion of the results for exclusions applied based on the clinical validity of the study population are provided below.

Patient Death in Episode: The mean observed cost for episodes ending in death was substantially higher than the mean observed cost for the final set of episodes: $3,404 compared to $1,895 at TIN level testing and $1,867 at TIN-NPI level testing. The mean O/E cost ratio for episodes ending in death is 1.83, which is much greater than the mean O/E cost ratio for final

[20] CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

episodes at both TIN-level testing (1.17) and TIN-NPI level testing (1.16). These results support the notion that this patient cohort is distinct in observed costs which is also reflected in O/E cost ratios. Excluding these episodes helps ensure that clinician performance is not inaccurately represented.

Not in OP, IP, or ASC Setting: Episodes not triggered in an IP, OP, or ASC setting are meant to be excluded, as episodes of care originating from other places of service may have different costs. This is corroborated by empirical data, where the mean O/E ratio is lower than the final set of reportable measures (0.85, compared to 1.17 and 1.16 at the TIN and TIN-NPI levels, respectively), while the mean, 50th, 75th, and 90th percentile grouped cost values are up to 71.5% greater than the same statistics on the final set of reportable episodes at both the TIN and TIN-NPI levels.

Mohs Surgery: Episodes including Mohs surgery with a melanoma diagnosis are excluded, as Mohs surgery is not part of the intended scope of the Melanoma Resection measure treatment arc, which is focused on standard surgical incisions for excising melanomas. These episodes have different cost statistics than the final reportable episodes, providing evidence that these episodes are different than the final episode group. The mean O/E ratio for episodes containing a Mohs surgery procedure (1.09) is lower than the mean O/E ratio for final reportable episodes at both the TIN (1.17) and TIN-NPI (1.17) levels. Furthermore, the 10th ($322) and 90th ($3,310) percentile observed costs are notably different for Mohs surgery episodes compared to Melanoma Resection episodes at the TIN (10th: $434; 90th: $4,237) and TIN-NPI (10th: $449; 90th: $4,199) level. These results support the clinical rationale for excluding these episodes.

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 154 risk factors and stratification by 2 risk categories.

The risk adjustment model for the Melanoma Resection 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. Patient age is included as one of 12 age categorical variables derived from the MA risk adjustment model's age/sex 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 patient'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 patients without a full 120-day lookback period are excluded from the measure. This 120-day period is used to measure patient health status and ensures that each patient'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 patient qualifies for Medicare through Disability or ESRD. The model also includes an indicator of whether the patient recently required long-term care, defined as 90 days in a long-term care facility without being discharged to community for 14 days. Patients who need to reside in long-term care facilities typically require more intensive care than patients 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 trigger procedure was performed in:
    - An ASC, off-campus outpatient hospital, or office/clinic
    - An on-campus outpatient hospital, to account for additional patient comorbidities and potentially more severe disease compared to excisions performed in an office/clinic
    - An inpatient place of service as a procedure performed in an IP stay can indicate a patient with comorbidities
- Whether the trigger excision procedure was performed:
    - on the ear or external ear canal to account for surgical complexity and downstream reconstruction due to body location
    - on the eyelid to account for surgical complexity and downstream reconstruction due to body location
    - on the lip, excluding the Vermilion border, to account for the additional clinical complexity in the excision procedure due to the location on the body and for the difficulty in repairing the surgical site factoring in aesthetic concerns
    - on the nose to account for the complexity in the excision procedure due to the location on the body and for the difficulty in repairing the surgical site when factoring in aesthetic concerns
- Whether the trigger excision procedure had an incision:
    - Greater than 4 centimeters to account for the inherently higher cost of excising larger melanomas to attain local disease control, or
    - Less than or equal to 4 centimeters to account for differences in excision size due to the size/location of the index melanoma
- In a teaching hospital where surgery was performed by a resident to account for potential additional complications due to inexperience
- Whether a flap or graft reconstructive procedure was performed in the post-trigger period to account for more complex melanomas requiring additional follow-up to close surgical sites and/or remedy aesthetic concerns
- Whether the patient had received services that would render them immunosuppressed in the 120 days prior to the trigger procedure which could impact wound healing, leaving patients more susceptible to downstream complications
- Whether the trigger procedure involved a melanoma in situ to account for the less severe and invasive nature of the resection procedure compared to larger melanomas
- Whether the patient underwent a tissue transfer/rearrangement:
    - Greater than 30 cm to account for the inherently higher cost to repair tissue after removing sufficient epidermal tissue to achieve sufficiently wide excision margins to attain local disease control
    - Less than or equal to 30 cm in the 90-day post trigger period to account for the additional cost of repairing a surgical site required by the trigger excision to achieve sufficiently wide margins for local disease control

- Whether the patient underwent a sentinel lymph node biopsy post-trigger to account for higher cost due to a potentially more severe disease state
- Whether the patient had received systemic chemotherapy or immunotherapy prior to the trigger procedure to account for potentially diminished immune response that could increase the risk for downstream complications, such as infection

Episodes with the highest 2% of observed costs are excluded before risk adjustment to mitigate the impact that these ultra-high-cost episodes could have on clinician measure scores, as these would generate very large O/E ratios that could dominate a clinician's O/E distribution.

The risk adjustment for this measure uses a log-linear regression model. The log-linear model is preferable to linear model as the melanoma episode cost is better characterized by a log-normal distribution than a normal distribution. The predicted, or expected, value is output on a log-scale by the risk adjustment model, which is then put through an exponential function with variance adjustment to be converted back to a standard cost scale for use in determining O/E ratios.

Finally, the risk adjustment model outlined above is stratified for each of the 2 Melanoma Resection measure sub-groups below, which are based on the body location of the melanoma being resected.

- Head/Neck
- Trunk/Extremity

Full details of the risk adjustment model are in the draft Measure Codes List File available on the MACRA Feedback page.[21]

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 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 Melanoma Resection measure methodology.

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 of 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 sub-groups, which may qualify as "ordering" of risk factors. 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. The sub-groups are listed in the above section. Melanoma excision procedures performed on the trunk or extremities versus on the head or neck were split into sub-groups to group patients into more clinically homogenous groups. Per expert clinical input, melanomas on the head and neck are clinically distinct from melanomas on the trunk and extremities, as they tend to be more clinically complex

[21] CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

and have more substantial aesthetic considerations for post-trigger complications and surgical site repair.

3.4.3 Conceptual Model of Impact of Social Risks

Our conceptual model of the impact of social risk factors is informed by both published external research and our own data analysis.[22,23,24]

3.4.4 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., Accountable Care Organizations, previous physician Quality and Resource Use Report programs, and other measures such as NQF #3512: Knee Arthroplasty, NQ #3509: Routine Cataract Removal with Intraocular Lens (IOL) Implantation, NQF #3510: Screening/Surveillance Colonoscopy, and 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 Evaluation of the CMS-HCC Risk-Adjustment Model report[25] and the Report to Congress: Risk Adjustment in Medicare Advantage[26]. 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

Acumen analyzed gender, dual status, income, education, and unemployment as social risk factors (more information on these variables can be found in Section 3.1.7). Patient 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. Patients without geographic information necessary to obtain ACS data were excluded, representing approximately 1.2% of episodes.

The percentage of male beneficiaries is much higher than the percentage of female beneficiaries, ranging from 57.98% (Trunk/Extremity) to 75.81% (Head/Neck), compared to 24.19% (Head/Neck) to 42.04% (Trunk/Extremity) across both sub-groups in this measure. This is corroborated by findings from the American Academy of Dermatology, who indicate that men are twice as likely at age 65 and 3 times as likely at age 80 to develop melanoma compared to women.[27] Regarding the different gender-specific rates of melanoma between sub-groups, one study noted that women are more prone to lower-extremity melanomas due to patterns of skin

[22] 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.
[23] 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
[24] 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/.
[25] Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.
[26] CMS, "Report to Congress: Risk Adjustment in Medicare Advantage," https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf.
[27] "Melanoma Strikes Men Harder." American Academy of Dermatology. Accessed July 8, 2020. https://www.aad.org/public/diseases/skin-cancer/types/common/melanoma/men-50.

exposure.[28] The vast majority of the beneficiaries (96.6 to 96.8%) 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% of observations. While 0.61 to 0.79% of beneficiaries are classified below a high school education level, the overwhelming majority (91.24 to 92.25%) of beneficiaries are classified at a high school level or greater. Finally, 12.30 to 12.81% 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 adjustors. Acumen ran a step-wise regression to include the following additional social risk factors on top of the adapted CMS-HCC model:

- Gender
- Dual status
- Gender + dual status
- Gender + dual status + race
- Gender + dual status + income + education + unemployment
- Gender + dual status + AHRQ SES index score
- Gender + dual status + race + income + education + unemployment
- Gender + dual status + race + AHRQ SES index score

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 patients for the relevant characteristic. However, the analysis also shows that the significance of the effects of social risk factors is not consistent. For example, female gender is statistically significant in the Head/Neck Melanoma sub-group, but is not statistically significant in the Trunk/Extremity sub-group in all risk adjustment models tested. Additionally, Age: 95+ is negatively correlated with risk-adjusted cost with a p-value of 0.00 in the Head/Neck sub-group, but had a p-value of 0.80 in the Trunk/Extremity sub-group.

Second, we analyzed the impact of adding social risk variables on overall model performance by looking at the differences in the ratio of observed to expected episode cost (O/E) 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 clinicians 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 measure scores greater than 98.6% of TINs and TIN-NPIs did not change or changed by ±5 percentiles or less.

[28] Stanienda-Sokół, Karolina, Natalia Salwowska, Martyna Sławińska, Katarzyna Wicherska-Pawłowska, Anna Lorenc, Dominika Wcisło-Dziadecka, Jerzy Wydmański, and Wojciech Majewski. “Primary Locations of Malignant Melanoma Lesions Depending on Patients’ Gender and Age.” Asian Pacific journal of cancer prevention : APJCP. West Asia Organization for Cancer Prevention, November 26, 2017. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5773794/.

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 level, with a Spearman correlation coefficient of 0.9988, and the TIN-NPI level with a correlation coefficient of 0.9987. These results indicate that the inclusion of social risk factors in the current risk adjustment model would have a limited effect on measure scores.

Based on these results, we believe the Melanoma Resection measure risk adjustment model sufficiently accounts for the effects of social risk factors on clinician measure scores.

3.4.6 Method for Statistical Model or Stratification Development

To analyze the validity of current risk adjustment model, we examined 2 analyses: (1) R-squared and adjusted R-squared for the regression models, and (2) predictive ratios and O/E cost ratios to examine the fit of the models at different levels of patient complexity.

1) R-squared and adjusted R-squared were calculated for the measure. These results should be evaluated in the context of the measure's service assignment rules which are intended to ensure only clinically associated costs are grouped to episodes. This is an important distinction from all-cost measures as service assignment leaves less variation for the risk adjustment model to explain. In this context, a low R-squared may indicate the effectiveness of the service assignment rules. These results are provided in Section 3.4.7.
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 3.4.8 and 3.4.9.

3.4.7 Statistical Risk Model Discrimination Statistics

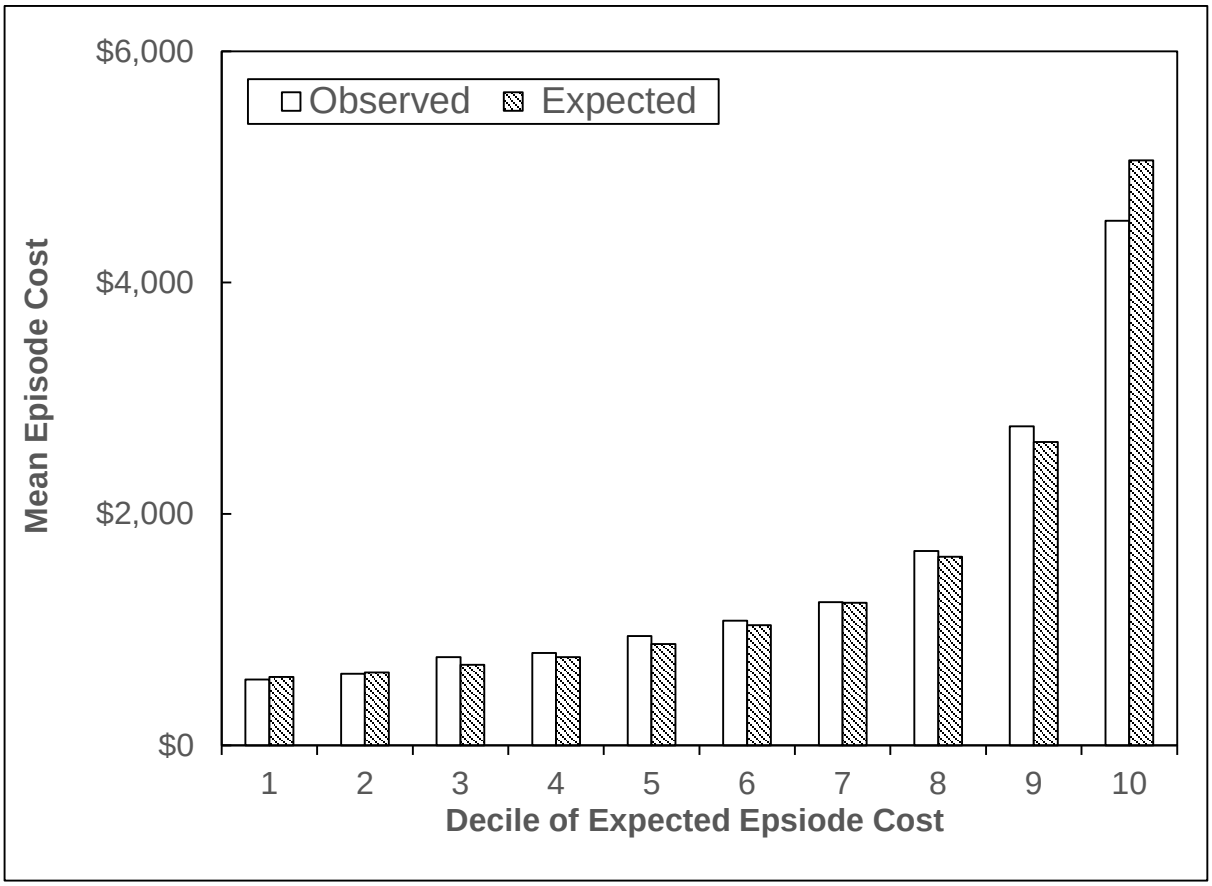
The overall R-squared for the Melanoma Resection cost measure, calculated by dividing explained sum of squares by total sum of squares is 0.56. The adjusted R-squared is 0.56. More information on discrimination testing for the CMS-HCC model can be found at Pope et al. 2011.[29] Note that the R-squared for the measure is for the log-linear model, which is not directly comparable to that of a linear model.

3.4.8 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 ratio for each risk decile varies from 0.93 in the highest risk decile to 1.10 in the third risk decile with an average O/E of 1.02 for all episodes.

Figure 1. Risk Adjustment Model Diagnostics: Comparison of Observed and Expected Cost by Predicted Cost Risk Deciles

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



3.4.9 Statistical Risk Model Calibration – Risk Decile

Analysis of predictive ratios by risk decile for the measure shows moderate variation among risk deciles, as predictive ratios range from 0.91 to 1.12 across all risk deciles (with an overall average of 1.01). Excluding the highest risk decile (with a predictive ratio of 1.12) results in a much narrower range of 0.91 to 1.04.

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.[30] As noted in Section 3.4.6 and 3.4.7, these results should be interpreted alongside service assignment rules, which remove clinically unrelated services, as well as the log-linear model, which transforms dollar figure to log scale, resulting in much smaller total sum of squared.

As demonstrated in Section 3.4.8 and 3.4.9, average O/E cost ratios and predictive ratios have moderate variation but are centered around one. Generally, we believe that the observed variation in these statistics may be due to inconsistent prediction from the log-linear risk adjustment model. We are currently evaluating the appropriateness of the log-linear risk adjustment model within the Melanoma Resection measure, and are exploring methods to account for the over/under correction of the model for further refinement of the Melanoma Resection measure.

[30] 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.

3.5 Identification of Meaningful Differences in Performance

3.5.1 Method

Our method of determining clinically meaningful differences in episode-based cost measure performance consists of stratifying clinician measure O/E cost ratios by meaningful characteristics and investigating the clinician O/E cost ratio distribution by percentile. The cost measure score numerator is the sum of the O/E cost ratio for all episodes attributed to a clinician. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. The denominator is the total number of episodes from the attributed to a clinician. Using O/E cost ratios allows for direct comparisons of performance at the sub-group level since a dollar figure cannot be calculated for those episodes using the national average observed episode cost. 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 or clinician group. We analyze the distribution of measure O/E cost ratios for clinicians defined by these characteristics.

The purpose of this analysis is to ensure that there is a sufficiently large difference in measure O/E cost ratios 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.

3.5.2 Statistical Results

Key findings show that, generally, there is a large performance difference among clinicians in the Melanoma Resection measure:

(i) The 99th percentile of the measure O/E cost ratio is nearly quadruple the measure O/E cost ratio at the 1st percentile for both the TIN level (0.54 to 1.99) and TIN-NPI (0.54 to 2.04) levels; and
(ii) The Melanoma Resection measure O/E cost ratio at the 90th percentile is approximately 90% greater than the O/E cost ratio at the 10th percentile at the TIN level (0.72 to 1.36) and 100% greater at the TIN-NPI level (0.70 and 1.40).

These results indicate there is a large potential for saving Medicare spending.

The results also show that there is not a systemic regional difference in clinician O/E cost ratios. For instance, the mean measure O/E cost ratios for clinicians across 9 census divisions (excluding 'Unknown') are within a range of 0.1 (i.e., 0.98 to 1.05 at the TIN level and 0.95 to 1.05 at the TIN-NPI level). Similarly, provider performance variation between urban and rural providers at both the TIN and TIN-NPI level are within a range of 0.08 (i.e., 1.01 for urban and 0.98 for rural at the TIN level, and 1.02 for urban and 0.94 rural providers at the TIN-NPI level). While there is some variation at the TIN-NPI level, this could be due to lower provider counts in rural areas, where less than 10% of TIN-NPIs involved in the measure are classified as rural.

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 procedures. The mean measure O/E cost ratio by number of episodes at a TIN level varies within a range of 0.07 (0.98 to 1.05). The mean measure O/E cost ratio by number of episodes at the TIN-NPI level does show some moderate variation, ranging from 0.99 to 1.11. However, the most extreme measure O/E cost ratios (i.e., furthest from 1.00) amongst the episode volume categories have the lowest provider counts, which may account for anomalous average scores.

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 riskier patients. Measure O/E cost ratios show some variation by risk score decile, with a range

in median TIN O/E cost ratio of 0.88 to 1.03 and a range in median TIN-NPI O/E cost ratio of 0.87 to 1.03.The lowest values in these ranges appear in the lowest 2 risk deciles at both the TIN and TIN-NPI levels, and the highest values appear between the seventh and ninth-highest risk deciles at the TIN and TIN-NPI levels.

Tables 5-A and 5-B below present the distribution of cost measure O/E cost ratios by a range of clinician/clinician group characteristics, allowing a comparison of O/E cost ratio distributions for these breakdowns. The cost measure O/E cost ratios are presented at the TIN (Table 5-A) and TIN-NPI (Table 5-B) level.

Table 5-A: Melanoma Resection TIN Level Cost Measure O/E Ratios

Characteristic	# of TINs	Mean O/E Ratio	O/E Ratio Percentile						
			1st	10th	25th	50th	75th	90th	99th
All TINs	1,812	1.01	0.54	0.72	0.83	0.95	1.13	1.36	1.99
Sub-group									
Head/Neck	1,730	0.96	0.34	0.58	0.70	0.88	1.11	1.40	2.49
Trunk/Extremity	1,794	1.03	0.51	0.71	0.83	0.95	1.13	1.39	2.42
Urban/Rural									
Urban	1,624	1.01	0.54	0.72	0.83	0.96	1.13	1.35	2.00
Rural	182	0.98	0.59	0.67	0.78	0.91	1.11	1.40	1.97
Unknown	2	1.13	0.97	0.97	0.97	1.13	1.28	1.28	1.28
Census Region									
Northeast	300	0.99	0.55	0.72	0.83	0.95	1.11	1.33	1.92
Midwest	311	1.02	0.54	0.73	0.83	0.96	1.15	1.41	2.00
South	801	1.01	0.53	0.72	0.82	0.95	1.12	1.35	2.13
West	393	1.02	0.56	0.72	0.85	0.96	1.13	1.35	2.03
Unknown	7	1.09	0.75	0.75	0.78	1.10	1.28	1.44	1.44
Census Division									
New England	98	1.01	0.50	0.74	0.84	0.97	1.11	1.33	1.97
Middle Atlantic	202	0.98	0.63	0.72	0.82	0.93	1.11	1.34	1.83
East North Central	190	1.01	0.54	0.72	0.83	0.95	1.14	1.40	1.80
West North Central	121	1.05	0.54	0.74	0.83	0.97	1.16	1.47	2.07
South Atlantic	521	1.01	0.54	0.72	0.84	0.96	1.12	1.36	2.20
East South Central	117	0.98	0.58	0.70	0.81	0.93	1.09	1.27	1.87
West South Central	163	1.01	0.44	0.69	0.81	0.94	1.17	1.32	2.13
Mountain	160	1.00	0.56	0.69	0.83	0.93	1.10	1.37	2.03
Pacific	233	1.03	0.62	0.74	0.86	0.98	1.15	1.34	1.95
Unknown	7	1.09	0.75	0.75	0.78	1.10	1.28	1.44	1.44
TIN risk score decile									
1st	181	0.96	0.51	0.69	0.76	0.89	1.03	1.23	2.59
2nd	181	0.91	0.36	0.66	0.77	0.88	1.01	1.22	1.58
3rd	181	0.98	0.44	0.71	0.82	0.93	1.08	1.29	1.94
4th	182	0.98	0.56	0.71	0.83	0.96	1.12	1.25	1.80
5th	181	1.03	0.53	0.75	0.84	0.96	1.16	1.34	2.24
6th	181	1.06	0.55	0.72	0.87	1.00	1.21	1.42	2.13
7th	182	1.06	0.54	0.73	0.83	0.96	1.21	1.44	2.20
8th	181	1.10	0.59	0.71	0.86	1.03	1.25	1.57	2.20
9th	181	1.02	0.61	0.75	0.86	0.96	1.13	1.34	1.90
10th	181	1.01	0.61	0.77	0.90	0.97	1.12	1.23	1.64
Number of episodes									
10-19 Episodes	823	1.00	0.49	0.68	0.78	0.94	1.13	1.44	2.20
20-39 Episodes	556	1.00	0.57	0.73	0.84	0.95	1.11	1.30	2.04
40-59 Episodes	193	1.02	0.65	0.75	0.88	0.99	1.13	1.29	1.84
60-79 Episodes	89	1.04	0.68	0.85	0.91	0.98	1.15	1.30	1.79
80-99 Episodes	48	1.00	0.67	0.82	0.89	0.95	1.07	1.19	1.49
100-199 Episodes	82	1.05	0.73	0.84	0.90	1.00	1.15	1.34	1.78
200-299 Episodes	11	0.98	0.85	0.88	0.91	0.94	1.06	1.13	1.16
300+ Episodes	10	0.99	0.89	0.89	0.90	0.98	1.07	1.13	1.15

Table 5-B: Melanoma Resection TIN-NPI Cost Measure O/E Ratios

Characteristic	# of TIN-NPIs	Mean O/E Ratio	O/E Ratio Percentile						
			1st	10th	25th	50th	75th	90th	99th
All TINs	2,188	1.01	0.54	0.70	0.82	0.96	1.14	1.40	2.04
Sub-group									
Head/Neck	2,025	0.96	0.34	0.56	0.68	0.87	1.11	1.45	2.54
Trunk/Extremity	2,145	1.03	0.50	0.70	0.81	0.96	1.14	1.41	2.42
Urban/Rural									
Urban	1,979	1.02	0.54	0.71	0.83	0.96	1.15	1.41	2.04
Rural	187	0.94	0.54	0.65	0.74	0.87	1.04	1.26	2.14
Unknown	1	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
Census Region									
Northeast	359	0.99	0.50	0.70	0.82	0.95	1.13	1.34	1.90
Midwest	368	1.03	0.53	0.72	0.82	0.97	1.17	1.44	2.02
South	986	1.01	0.54	0.71	0.82	0.94	1.13	1.40	2.20
West	452	1.02	0.56	0.69	0.82	0.97	1.13	1.41	2.14
Unknown	23	1.02	0.54	0.68	0.84	0.98	1.28	1.48	1.60
Census Division									
New England	134	1.04	0.50	0.74	0.88	0.99	1.18	1.42	1.92
Middle Atlantic	225	0.96	0.52	0.68	0.80	0.92	1.10	1.29	1.68
East North Central	243	1.02	0.52	0.71	0.82	0.98	1.15	1.41	1.91
West North Central	125	1.05	0.56	0.75	0.83	0.96	1.18	1.48	2.08
South Atlantic	676	1.02	0.56	0.72	0.83	0.96	1.15	1.40	2.22
East South Central	124	0.95	0.54	0.66	0.76	0.92	1.10	1.27	1.87
West South Central	186	1.01	0.44	0.69	0.81	0.93	1.10	1.41	2.22
Mountain	193	0.99	0.55	0.69	0.78	0.95	1.09	1.29	2.14
Pacific	259	1.04	0.56	0.70	0.83	1.00	1.17	1.44	2.22
Unknown	23	1.02	0.54	0.68	0.84	0.98	1.28	1.48	1.60
TIN-NPI risk score decile									
1st	218	0.93	0.60	0.68	0.75	0.87	1.01	1.18	1.91
2nd	219	0.95	0.49	0.66	0.77	0.89	1.07	1.27	1.83
3rd	219	0.96	0.43	0.66	0.73	0.89	1.05	1.37	2.02
4th	219	0.99	0.55	0.68	0.80	0.95	1.10	1.32	2.01
5th	219	1.02	0.53	0.70	0.85	0.97	1.16	1.39	2.03
6th	219	1.07	0.54	0.68	0.82	0.99	1.23	1.54	2.22
7th	219	1.09	0.51	0.73	0.86	1.03	1.25	1.52	2.28
8th	219	1.07	0.57	0.76	0.84	1.01	1.23	1.46	2.00
9th	219	1.08	0.58	0.75	0.89	1.03	1.20	1.48	1.97
10th	218	0.98	0.63	0.79	0.87	0.97	1.08	1.18	1.46
Number of episodes									
10-19 Episodes	1,371	0.99	0.51	0.68	0.79	0.93	1.11	1.40	2.02
20-39 Episodes	626	1.05	0.63	0.75	0.85	0.99	1.16	1.43	2.15
40-59 Episodes	137	1.07	0.59	0.72	0.92	1.04	1.18	1.39	2.08
60-79 Episodes	33	1.11	0.67	0.82	0.97	1.05	1.22	1.41	1.97
80-99 Episodes	14	1.05	0.74	0.80	0.88	1.02	1.19	1.43	1.52
100-199 Episodes	7	1.04	0.84	0.84	0.87	0.89	1.28	1.32	1.32
200-299 Episodes	0	-	-	-	-	-	-	-	-
300+ Episodes	0	-	-	-	-	-	-	-	-

3.5.3 Interpretation

The results in Tables 5-A and 5-B above indicate that there is limited overall variation in the mean cost measure O/E cost ratio across episode sub-groups, the urban/rural divide, census

regions, and census divisions at both the TIN and TIN-NPI levels. For each of these variables, the difference in the mean O/E cost ratio across categories was 0.1 or less. This indicates that the risk adjustment model is overall functioning as intended; it is adjusting cost performance such that there are no substantive differences across the categories for these variables.

As for sub-groups, the risk adjustment model is run individually for each sub-group to account for the different clinical considerations for excising and repairing head/neck melanomas compared to trunk/extremity melanomas and enable a more fair comparison across episodes. These results support that both sub-groups are able to capture meaningful differences in cost performance, even after risk adjustment. For each sub-group, there is at least a five-fold increase between the 1st and 99th percentile measure O/E cost ratios. These results indicate that there are substantial savings opportunities for Medicare and that there are no systemic differences across provider type, sub-group, or geographic provider characteristics.

For provider risk score decile, the difference in mean O/E cost ratio across categories was 0.19 at the TIN and 0.16 at the TIN-NPI level. Similar to the discussion in Sections 3.4.9 and 3.4.10, the variation in mean O/E cost ratio may indicate that the current risk-adjustment model may be overcorrecting for patients in lower risk deciles. Based on these and risk-adjustment model testing results, we are currently evaluating the appropriateness of the log-linear risk adjustment model within the Melanoma Resection measure, and are exploring methods to account for the over/under correction of the model for further refinement of the Melanoma Resection measure.

Regarding case volume, while mean TIN measure O/E cost ratios show no notable variation by case volume, there is moderate variation in provider O/E cost ratios at the TIN-NPI level, with a range of 0.12. However, as discussed in Section 3.5.2, this variation may be the result of low provider counts in the larger episode volume categories, as these deciles showed the most extreme mean O/E cost ratio values.

3.6 Missing Data Analysis and Minimizing Bias

3.6.1 Method

Since CMS uses Medicare claims data to calculate the Melanoma Resection measure, Acumen expects a high degree of data completeness. To further ensure that we have complete and accurate data for each patient who opens an episode, 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 death date occurs before the episode trigger date.

The Melanoma Resection 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 four categories of missing data which caused episodes to be excluded from the Melanoma Resection 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:

- Episode does not have a main surgeon
- Patient death date occurred before the trigger date

- Patient has a primary payer other than Medicare during the episode window or in the 120-day lookback period
- Patient was not enrolled in Medicare Parts A and B, or was enrolled in Part C, during the 120-day lookback period and episode window

As a note, the episode and clinician counts below reflect exclusion from the initial population of triggered episodes, which consists of 990,735 episodes across 797,449 beneficiaries that represent occurrences of a cutaneous excision or tissue rearrangement procedure. After the missing data exclusions are applied, we then apply additional trigger logic to this overall patient cohort to narrow the population to only episodes with a diagnosis of melanoma using the relevant diagnosis codes. After applying this additional trigger logic, there are 83,341 episodes for 70,179 patients. Additional information regarding the trigger logic can be found in Section 3.1.6.

Table 6: Missing Data Categories for the Melanoma Resection Measure

Exclusion	# Episodes	# TINs	# TIN-NPIs
No main surgeon	62	45	52
Death before trigger	*	*	*
Other primary payer	68,482	8,613	17,658
Not continuously enrolled	30,398	6,462	12,124

* indicates that there were fewer than 11 episodes

3.6.3 Interpretation

As the Melanoma Resection 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.

Contact Information

Other Additional Information
Melanoma Resection Clinician Expert Workgroup Members:
Aamir Siddiqui, MD, American Society of Plastic Surgeons
Anna Likhacheva, MD, MPH, American Society for Radiation Oncology
Clifford Lober, MD, JD, American Academy of Dermatology
Howard Rogers, MD, PhD, American College of Mohs Surgery
Jennifer Stein, MD, American Academy of Dermatology
Melissa Piliang, MD, American Academy of Dermatology
Michele Manahan, MD, MBA, FACS, American Society of Plastic Surgeons
Nita Kohli, MD, MPH, American College of Mohs Surgery
Oliver Wisco, DO, American Academy of Dermatology
Philip Devlin, MD, FACR, FASTRO, FFRRCSI, FABS, American Society for Radiation Oncology
Samir Khariwala, MD, MS, American Academy of Dermatology
Scott Collins, MD, American Academy of Dermatology
Victoria Lazareth, MA, MSN, NP-C, DCNP, American Association of Nurse Practitioners

The Melanoma Resection Clinician Expert Workgroup is composed from the larger Dermatologic Disease Management Clinical Subcommittee. The composition list of the Clinical Subcommittee is included in the Episode-Based Cost Measures Development Process document.[31]

[31] CMS, "2020 Episode-Based Cost Measure Field Testing Wave 3 Measure Development Process," MACRA Feedback Page, https://www.cms.gov/files/document/macra-cmft-ebcm-process-2020.pdf.