Evaluation of the Oncology Care Model:

Performance Periods 1–6 Appendices

December 2021

Contract #HHSM-500-2014-000261 T0003

SUBMITTED BY:

Abt Associates
Andrea Hassol, Project Director
6130 Executive Boulevard
Rockville, MD 20852

IN PARTNERSHIP WITH
The Lewin Group
Harvard Medical School
GDIT
Geisel School of Medicine at Dartmouth

PREPARED FOR:

Jessica McNeely
Center for Medicare & Medicaid Innovation
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244

Abt
ASSOCIATES

AUTHORS:

Abt Associates: Andrea Hassol, Nathan West

The Lewin Group: Carol Simon, Shalini Jhatakia, Inna Cintina, Yvette Overton, Amaka Ume, Maya Nilkant

Harvard Medical School: Nancy Keating, Mary Beth Landrum

Geisel School of Medicine at Dartmouth: Gabriel Brooks

General Dynamics Information Technology: Colleen Kummet, Van Doren Hsu, Stephanie Shao

ACKNOWLEDGMENTS:

BOLD
THINKERS
DRIVING
REAL-WORLD
IMPACT

Abt Associates | 6130 Executive Boulevard | Rockville, MD 20852

CONTENTS

A. Methods

A.1. Data and Methods for Analysis of Medicare Claims and Administrative Data

This appendix section contains information about the data and methods we used to construct payment outcome measures from Medicare claims for the Oncology Care Model (OCM) evaluation. The primary data sources used to measure OCM impacts on payment outcomes include the Common Medicare Environment and Enrollment Database files, 100 percent of the Medicare Parts A and B claims files, and 100 percent of the Part D Prescription Drug Event (PDE) files.

This appendix describes how claims and other data sources were used to construct outcome measures, the performance periods (PPs) included in this report, how episodes were identified for analysis, how the comparison group was constructed and validated, and the analytic approaches used to quantify impacts of the Model.

A.1.1 Secondary Data Sources

The data sources and how they were used to construct the analytic files are summarized in **Exhibit A-1**.

Exhibit A-1: Data Sources Used in the Claims Analysis

Data Source	Purpose
2014–2019 Part B Claims (Virtual Research Data Center, (VRDC))	• Identify Part B chemotherapy episode triggers for episode identification, and cancer-related evaluation and management (E&M) services for episode attribution. • Determine the presence of cancer diagnosis within 59 days prior to and including the service date of a Part D chemotherapy claim to identify Part D chemotherapy episodes. • Identify cancer-related E&M services from carrier claims during episodes. • Calculate episode-level payment measures for Part B services. • Construct Hierarchical Condition Category (HCC) scores.
2014–2019 PDE Tap Files (VRDC)	• Identify Part D chemotherapy triggers for episode identification. • Calculate episode-level Part D drug payment measures.
2014–2019 Part A Claims (VRDC)	• Calculate episode-level payment measures for Part A services. • Construct HCC scores.
2014–2019 Integrated Data Repository System	• Determine standardized Part A and B payments.

2014–2019 Common Medicare Environment Master Beneficiary Summary Files (VRDC)	• Determine Part A and B enrollment for beneficiary eligibility criteria for episode identification.
	• Determine:
	  ▪ Beneficiary characteristics including age, race, and gender
	  ▪ Beneficiary ZIP code of residence
	  ▪ Monthly Part D enrollment and dual eligibility
	  ▪ County-level Medicare Advantage Penetration
	  ▪ County-level emergency department visits among fee-for-service (FFS) population

2014–2019 Enrollment Database Files (VRDC)	• Determine Medicare Secondary Payer information for beneficiary eligibility criteria for episode identification.
2014–2019 Common Medicare Environment Files (VRDC)	• Determine end-stage renal disease coverage for episode identification.
2016–2019 Food and Drug Administration National Drug Code Directory	• Identify PDEs that are for drugs, excluding vaccines.

Data Source	Purpose
2016–2019 Medicare Part B Drug Average Sales Price	• Identify Part B claims that are indicative of drugs.
2014–2018 Centers for Medicare & Medicaid Services (CMS) Health Professional Shortage Area (HPSA) Files	• Identify proportion of the population within a county residing in an HPSA.
2014–2018 National Plan and Provider Enumeration System (NPPES)	• Supplement provider specialty information in Part B Claims data.
2014–2018 Master Data Management Beneficiary Extracts (VRDC)	• Identify beneficiary alignment to the following CMS initiatives: Pioneer Accountable Care Organization (ACO), Medicare Shared Savings Program, Next Generation ACO, Comprehensive Primary Care (CPC), and CPC Plus.
July 2015, August 2016, August 2017, and August 2018 SK&A[1] Office-Based Physician File	• Link practice sites to Tax Identification Numbers (TINs) to construct practice's affiliation with health system and hospital ownership.
2014–2019 Area Health Resource Files	• Construct county-level sociodemographic and market supply characteristics.
Welch and Bindman 2016,[2] list of Association of American Medical Colleges (AAMC) medical schools,[3] and websites of medical school oncology/hematology departments, divisions, and institutes	• Identify TINs that are affiliated with a medical school's academic medical group.
OCM program data	• Identify OCM practice participation. • Identify legacy TINs for OCM practices in baseline period. • Identify reconciliation-eligible episodes in each PP and associated expenditures. • Identify total amount paid by Medicare for performance-based payment (PBP) and Monthly Enhanced Oncology Services (MEOS).

---

[1]   http://www.skainfo.com/databases/physician-data

[2]   Welch, P and Bindman, AB. Town and gown differences among the largest medical groups in the US. *Journal of Academic Medicine* July 2016;91(7):1007–14.

[3]   AAMC Medical School Members. Available at:
https://members.aamc.org/eweb/DynamicPage.aspx?site=AAMC&webcode=AAMCOrgSearchResult&orgtype=Medical%20School

The Medicare claims used in this report were retrieved as of April 2020, and three months of claims run-out was applied uniformly. A report on Medicare claims maturity[4] estimates that over 90 percent of Part A and B claims and PDEs are received within three months of service, and approximately 90 percent of Part B claims are finalized within three months. This timing does not apply to claims for the monthly per-beneficiary $160 MEOS payment that practices may bill for to cover the provision of enhanced services and care coordination.

A.1.2   Observation Period for This Report

OCM began July 1, 2016 and focuses on six-month episodes of care triggered by chemotherapy FFS Medicare beneficiaries with continuous Parts A and B enrollment. OCM is organized into six-month PPs, for which CMS retrospectively assesses the performance of participating practices and reconciles payments. The six-year Model has a total of 11 PPs. The first PP included episodes that started between July 1, 2016 and January 1, 2017 and ended by June 30, 2017. The last PP will include episodes starting between July 2, 2020 and January 1, 2021, all of which will end by June 30, 2022.

Exhibit A-2 summarizes the observation period for this report, which covers OCM impacts through PP6. The baseline period includes six-month episodes that began July 2, 2014 through January 1, 2016 and ended between January 1, 2015 and June 30, 2016. The intervention period covered in this report includes six-month episodes that began during the Model's first six PPs (PP1–PP6), between July 1, 2016 and July 1, 2019, and ended between December 31, 2016 and December 31, 2019. The baseline period began in July 2014 to align with the calendar start of the Model, which started in July 2016. This alignment by calendar month addresses seasonality in Part D payments,[5] which must be studied symmetrically in both time periods.

Practice applications to participate in OCM were due to CMS on June 30, 2015, and CMS notified practices of acceptance into the Model in April 2016. CMS anticipated that accepted practices would make changes in staffing, resources, and care delivery in preparation for Model start. As a result, we apply a "hold-out" period so that early anticipatory practice changes do not contaminate the baseline period. Specifically, we do not include in the baseline episodes that began between January 2, 2016 and June 30, 2016 and ended between July 1, 2016 and December 29, 2016. Episodes that began during this period ended early in the first PP, which would have contaminated the baseline and intervention periods.

---

[4]   Chronic Condition Data Warehouse. (2017). CCW white paper: Medicare claims maturity. October. Version 2.0. Available from https://www.ccwdata.org/web/guest/ccw-medicare-data-white-papers.

[5]   As a consequence of the Medicare Part D benefit structure, Medicare payments are not observed on individual PDE records until a beneficiary enters catastrophic coverage (unless the beneficiary qualifies for low-income subsidy). As a result, most beneficiaries will not have PDEs with positive Medicare payments recorded until entry into the catastrophic phase, which on average occurs later in the calendar year. Previous analyses showed that among the six-month episodes of care used in the OCM evaluation, episodes that begin during the third quarter of the year tend to have the highest Part D payments, on average.

Exhibit A-2: Observation Period for the Report Covering PP1–6

Period	Performance Period	Episodes Triggering	Episodes Ending	Time Periods Specified for Difference-in Differences (DID) Analyses
Baseline-3	-3	7/2/14–1/1/15	1/1/15–6/30/15	
Baseline-2	-2	1/2/15–7/1/15	7/1/15–12/31/15	Baseline period
Baseline-1	-1	7/2/15–1/1/16	1/1/16–6/30/16	
Hold-out	0	1/2/16–6/30/16	7/1/16–12/29/16	Hold-out period
PP 1	1	7/1/16–1/1/17	12/31/16–6/30/17	
PP2	2	1/2/17–7/1/17	7/1/17–12/31/17	
PP3	3	7/2/17–1/1/18	1/1/18–6/30/18	Intervention period for report covering PP1–6
PP4	4	1/2/18–7/1/18	7/1/18–12/31/18	
PP 5	5	7/2/18–1/1/19	1/1/19–6/30/19	
PP 6	6	1/2/19–7/1/19	7/1/19–12/31/19	
PP 7	7	7/2/19–1/1/20	1/1/20–6/30/20	
PP 8	8	1/2/20–7/1/20	7/1/220–12/31/20	Intervention periods for future evaluation reports
PP 9	9	7/2/20–1/1/21	1/1/21–6/30/21	
PP 10	10	1/2/21–7/1/21	7/1/21–12/31/21	
PP 11	11	7/2/21–1/1/22	1/1/22–6/30/22	

Notes: PP: performance period. DID: difference-in-differences.

A.1.3 Episode Identification

We identified all eligible cancer episodes nationwide that occurred during the baseline period, and, separately, during the intervention period, following the OCM methodology.[6] **Exhibit A-3** shows the number of episodes used in this report, for the OCM and comparison groups, for each period.

First, we identified a Part B or Part D chemotherapy trigger event, defined as the first date of a Part B chemotherapy drug claim or Part D chemotherapy drug claim with a corresponding Part B claim for cancer within 59 days of the Part D claim, in each PP, assuming this date is not included in a previous episode. Then, among beneficiaries with a trigger chemotherapy event, we used Part B carrier claims to determine whether the beneficiary had had at least one cancer-related E&M service during the six months following the chemotherapy trigger event, billed under a TIN that has at least one oncology

Exhibit A-3: Number of Episodes by PP

Period	Number of Episodes	
(Episodes Initiating)	OCM	Comparison Group
Baseline-3 (7/2/14–1/1/15)	113,552	134,074
Baseline-2 (1/2/15–7/1/15)	117,335	138,560
Baseline-1 (7/2/15–1/1/16)	114,994	132,971
Hold-Out Period (1/2/16–6/30/16)		
PP1 (7/1/16–1/1/17)	126,654	145,234
PP2 (1/2/17–7/1/17)	128,238	146,648
PP3 (7/2/17–1/1/18)	124,327	138,790
PP4 (1/2/18–7/1/18)	132,814	145,987
PP5 (7/2/18–1/1/19)	129,418	140,333
PP6 (1/2/19–7/1/19)	137,418	147,758
Total All Periods	1,124,750	1,270,355

Source: Medicare Claims 2014-2019.

---

[6]   RTI International. (2018). OCM performance-based payment methodology. Version 5.1. Prepared for the Centers for Medicare and Medicaid Services in partnership with Actuarial Research Corporation. Research Triangle Park, NC: RTI International; December 17. Available from https://innovation.cms.gov/initiatives/oncology-care/

provider (National Provider Identifier (NPI)).[7] Finally, we required that the beneficiary meet the additional OCM inclusion criteria during the entire episode: continuous Medicare Parts A and B enrollment; coverage under Medicare FFS (not Medicare HMO, Medicare Advantage, or the United Mine Workers of America program); Medicare as the primary payer; and no Medicare benefit due to end-stage renal disease. An episode could end earlier than six months following the trigger event only if the beneficiary died.

A.1.4   Attribution of Episodes to Practices

After identifying all eligible episodes, per the OCM attribution methodology, we assigned episodes to the practice that provided the plurality of cancer-related E&M services during the episode.[8] A practice is defined as a TIN with at least one oncology provider. A TIN is a billing unit for tax purposes, and it may or may not represent the structure of a physician group organization; some oncology groups use multiple TINs, and some oncology groups share a single TIN with a larger multi-specialty organization. For OCM, CMS requires that participating practices each use a single TIN, and that all clinicians in the practice submit oncology claims under that TIN. Participating OCM practices that experienced billing or business changes during the baseline or intervention period provided CMS with any "legacy" (i.e., older) TINs to capture billing for the entire practice. We used these legacy TINs to attribute episodes to OCM practices in the baseline period. Because legacy TINs are not available for groups not participating in OCM (i.e., comparison TINs used for this evaluation), we were unable to track such organizational changes, and instead attributed episodes to individual comparison TINs. We therefore define a comparison practice as a TIN with at least one oncology provider.

A.1.5   Sample of OCM and Comparison Practices

OCM practices volunteered to participate in the Model and may differ from non-OCM practices. In the first six PPs, we included 202 practices participating in OCM. In selecting a comparison group, we sought to identify non-OCM TINs that, as a group, were similar to the group of OCM practices in the period prior to CMS's announcement of OCM. Comparison practices were selected using propensity score matching (PSM). The objective of PSM is to identify a comparison group that is statistically similar to the treatment group, based on observable factors.

First, starting from the universe of non-participating physician practices, we identified a subset of practices that were relevant for OCM and eligible to participate in OCM based on Model rules. From this subset we used PSM to identify comparison practices based on patterns of billing for OCM services and similarity to OCM practices in terms of key practice, beneficiary, and market characteristics. The PSM yielded 534 practices for the comparison group. Detailed information about the comparison group selection and PSM methodology is provided in the ***Performance Period One Report***. The PP1–PP6 intervention period as a whole had 522 comparison practices with attributed episodes; this number declined to 460 practices with episodes in PP5. Some attrition was anticipated, and the comparison group was deliberately constructed to be large enough to accommodate a modest reduction in TINs and episodes over time. Attrition was due to a variety of reasons including practice closures, mergers with or acquisitions by other practices or hospitals, or that the TIN no longer had attributed episodes.

---

[7]   The requirement that a TIN have at least one oncology provider was applied to all baseline and intervention PPs.

[8]   RTI International. (2018). OCM performance-based payment methodology. Version 5.1. Prepared for the Centers for Medicare and Medicaid Services in partnership with Actuarial Research Corporation. Research Triangle Park, NC: RTI International; December 17. Available from https://innovation.cms.gov/initiatives/oncology-care/.

A.1.6 Claims-Based Payment Outcome Measures

Exhibit A-4 defines each of the payment measures evaluated in this report.

Exhibit A-4: Definition of Medicare Payment Outcome Measures

Outcome Measure	Definition
Overall Payments	
Total episode payments (TEP) – Part A, B, and D payments	Total Part A, B, and D Medicare payments, not including MEOS payments, per episode. Part A and B payments are standardized. In other words, geographic differences in Medicare payment rates (e.g., due to variations in local wages or input prices) as well as payment variation resulting from CMS program reductions/additions (e.g., for programs including bundled payment), were removed. Part D payments are not standardized and were measured as the sum of low-income cost-sharing amount and 80 percent gross drug cost above the out-of-pocket threshold. All payments reflect the Medicare payment, not allowed payments.
Part A payments	Total Part A Medicare payments per episode. Part A spending includes payments for acute care hospitalizations, hospitalizations at other inpatient facilities, post-acute care (i.e., services at skilled nursing facilities, inpatient rehabilitation facilities, long-term care hospitals, and home health agencies), and hospice care.
Part B payments (without MEOS)	Total Part B Medicare payments, excluding MEOS payments, per episode. Part B payments include spending on hospital outpatient services, physician services, and durable medical equipment.
Part D payments	Total Part D Medicare payments per episode. Part D payments are typically for oral prescription drugs obtained at the pharmacy. This measure was restricted to episodes for beneficiaries enrolled in Part D for all months of the episode, while alive.
Part D gross drug costs	Total Part D gross drug costs per episode. A prescription's gross drug costs, reflects payments made by all parties (beneficiary, plan, Medicare), and is calculated as the sum of ingredient cost, dispensing fee, sales tax, and vaccine administration fee. This measure was restricted to episodes for beneficiaries enrolled in Part D for all months of the episode, while alive.
Part A Payments Components	
Acute care hospital (ACH) payments	Payments for ACH hospitalization(s) per episode. The measure includes ACH hospitalizations that originated during the episode (i.e., claim from date on the hospitalization occurred within the episode start and end dates). The full payment of the hospitalization was allocated to the episode, even if the hospitalization extended beyond the end of the episode.
Skilled nursing facility (SNF) payments	Payments for post-acute SNF stays per episode (claim type 20, 23). The full payment of the SNF stay was allocated to the episode, even if the stay extended beyond the end of the episode.
Home health agency payments	Payments for post-acute home health agency services per episode (claim type 10).
Hospice payments	Payments for hospice services per episode (claim type 50).
Inpatient rehabilitation facility payments	Payments for post-acute services at an inpatient rehabilitation facility per episode (claim types 60, 61).
Long-term care hospital payments	Payments for post-acute services at a long-term care hospital per episode (claim types 60, 61).

Outcome Measure	Definition
Part B Payments Components	
Imaging payments	Payments for standard, advanced, and other imaging services per episode. Standard and other imaging included x-ray, echography, and cardiac catheterization. Advanced imaging included computerized axial tomography scans, magnetic resonance imaging, and nuclear medicine (e.g., positron emission tomography).
Laboratory payments	Payments for laboratory services per episode.
E&M payments	Payments for E&M services per episode.
Chemotherapy, Cancer-Related, and Drug Payments	
Part B chemotherapy payments	Part B chemotherapy payments per episode. Part B chemotherapy drugs were identified using the Healthcare Common Procedure Coding System (**HCPCS**) codes found in the chemotherapy trigger list, per OCM Model specifications.
Part B non-chemotherapy drug payments	Payments for Part B non-chemotherapy drugs per episode.
Radiation therapy payments	Payments for Part B radiation therapy services per episode. Procedure codes for radiation therapy were identified per OCM Model specifications.
Cancer-related E&M payments	Payments for Part B cancer-related E&M services per episode. A cancer-related E&M service was defined as an E&M service in a non-institutional setting with a cancer diagnosis on the same line (per OCM Model specifications for episode identification and attribution).

A.1.7 Approach for Claims-Based Analyses

In this section, we describe the claims-based impact analyses conducted for this Annual Report. Analyses were conducted in CMS's VRDC environment using SAS Enterprise Guide v7.1 and Stata/MP 16.1 statistical software.

Impact Analyses

Given the quasi-experimental design of OCM, we use DID regression analyses to estimate Model impact on important payment outcomes. DID is a statistical technique that quantifies the impact of an intervention by comparing changes in outcomes of treatment cases (in this case, OCM episodes) to changes in outcomes in a matched comparison group (comparison episodes), from before to after Model implementation. The DID results describe the average effect of OCM over the entire duration of the intervention period, and for each of the first five PPs individually.

We performed all DID analyses at the episode level. We used ordinary least squares regression models for payment outcome measures. The models were specified to derive estimates of the impact of OCM for each PP quarter. Using a weighted average, [9] we then combined PP quarter estimates into a single cumulative impact estimate and individual PP estimates (two quarters per PP). Because multiple episodes were attributed to the same practice, provider patterns or actions that affect all episodes attributed to a practice will result in errors that are correlated. As a result, we adjusted standard errors to reflect the fact that episodes were clustered at the practice level. Our models also included state fixed effects to adjust for state-level characteristics (e.g., regulations, policies) not otherwise captured by the covariates included in the models (see below). [10]

DID Specification

The growth rate of many payment outcome measures varied considerably by cancer episode type, over time. For example, in PP4 and PP5, there was a sharp increase in TEP for lung cancer episodes that was not present in PP1 to PP3; in contrast, for colorectal cancer episodes, the change in TEP (relative to baseline) was the same in all PPs. These differences by cancer episode type were likely due to the availability of new, more expensive treatments used for specific cancer types in more-recent PPs. To account for these varied trajectories by cancer episode type, we incorporate cancer interactions in the DID specification used to assess payment measures. Including these interaction terms in the specification improved model fit.

The form of the DID specification we use for assessing payment outcomes is as follows:

$$Y = \beta_0 + \beta_1 OCM + \sum_{q=1}^{N} \gamma_q PPQ_q + \sum_{c=1}^{G} \partial_c Can_c + \sum_{q=1}^{N} \alpha_q OCM \cdot PPQ_q + \sum_{c=1}^{G} \theta_c OCM \cdot Can_c + \sum_{q=1}^{N} \left(\sum_{c=1}^{G} \delta_{qc} Can_c \cdot PPQ_q \right) + \sum_{q=1}^{N} \left(\sum_{c=1}^{G} \beta_{qc} OCM \cdot Can_c \cdot PPQ_q \right) + \beta' X + \varepsilon, \textbf{(1)}$$

where Y is an outcome for each episode originating in quarter q; OCM is an indicator distinguishing OCM practices from comparison practices; PPQ is an indicator distinguishing each quarter of intervention data from the baseline data; Can is an indicator distinguishing the 24 cancer types and the group of non-reconciliation-eligible cancer types; and X is a vector of pre-determined covariates for each episode. The indicators for OCM, PP quarter and cancer type are interacted to account for cancer-specific trajectories in payments and use between the baseline and intervention periods, as described above.

---

[9] Calculating cumulative and PP-level estimates from weighted quarterly averages accounts for changing distributions and number of episodes over time.

[10] State fixed effects were added to cancer type-specific models. State fixed effects were excluded from these models in previous reports due to sample size limitations.

The coefficient β_{qc} in model (1) captures the incremental, or marginal, impact of the OCM intervention on outcome **Y**, for cancer type **c**. The β_{qc} coefficients are aggregated across all cancer types to estimate the impact of OCM in each PP quarter, relative to changes over the same time period in episodes of comparison practices. We use the estimated coefficients to generate predicted values of the outcome measures. We compare two predictions to calculate the marginal effect. The marginal effect is equal to the average marginal effect for each observation, which is calculated as the difference between the predicted treatment outcome and a predicted counterfactual outcome where the impact of OCM is assumed to be zero.[11] Using this model, we constructed estimates of the overall impact of OCM, and impact for specific PPs by taking linear combinations of the estimates of the appropriate PP quarters. We weighted the PP quarter estimates by the number of episodes in each PP quarter to obtain the average cumulative and PP-level impacts, and used the delta method to assign significance to combined estimates.

In addition to the DID estimates, we present regression-adjusted means of the outcome measures for OCM and comparison episodes during the baseline and intervention periods, and examine trends across the two periods. We also present the DID estimate as a percentage of the OCM baseline mean to provide context (scale) and quantify the relative percentage change associated with OCM. Finally, for some key payment measures, we calculate trends reflecting the risk-adjusted mean in the outcome measure for each PP from the start of baseline until PP6, separately for OCM and comparison episodes.

Covariate Selection

The DID model controls for time-varying changes/influences that affect both the comparison and OCM groups, as long as model assumptions are met, **Exhibit A-5** shows the beneficiary-, practice-, and market-level factors we control for in DID analyses. The covariates in the DID models were informed by the broader research literature on oncology outcomes, a review of National Quality Forum measures,[12] discussions with clinical experts, and extensive statistical testing of alternative specifications using baseline period data. We include 27 covariates in all DID impact analyses. For a small group of outcomes, we exclude redundant covariates to achieve model convergence. For example, for all Part D-related outcome measures that apply to beneficiaries enrolled in Part D, the covariate indicating Part D enrollment is excluded.

Exhibit A-5: Covariates Included in DID Models

Domain	Model Covariate	Definition
Beneficiary-Level		
Beneficiary characteristics	Sex	Beneficiaries are categorized as male or female based on documented sex.
	Race/ethnicity	Beneficiaries are categorized as non-Hispanic White, Black (or African-American), Hispanic, or Other (Asian/Pacific Islander, American Indian, Other, Unknown) based on RTI race code methodology.
	Age	Beneficiaries are categorized as under 65, 65–69, 70–74, 75–79, 80–84, and 85+ years of age.
	Medicaid dual eligibility	Beneficiaries are categorized as having full/partial Medicaid benefits or having no benefits.
	Part D enrollee	Beneficiaries are coded as a Part D enrollee if enrolled in Part D for all months of the episode, while alive.

---

[11]   Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters* 115(1):85–87.

[12]   National Quality Form. (2018). National Quality Forum [Internet homepage]. [Updated March 23, 2003; cited November 9, 2003]. Available from http://www.qualityforum.org/Home.aspx.

Domain	Model Covariate	Definition
CMS program alignment	Beneficiary alignment to other CMS programs	Beneficiaries are coded as aligned if they were involved in at least one of the following CMS initiatives during their episode: Pioneer ACO, Medicare Shared Savings Program, Next Generation ACO, CPC, or CPC+.
Beneficiary clinical characteristics	Cancer type	The 24 cancer episode types of interest are derived from the cancer types assigned to each episode per the OCM methodology. Each episode is assigned a cancer type using the plurality of cancer diagnoses on E&M services in the carrier file that occurred during the episode. The 21 reconciliation-eligible cancer types in the original OCM methodology were expanded to 24, with breast cancer divided into low-risk and high-risk episodes, prostate cancer divided into low-intensity and high-intensity episodes,[13] and bladder cancer divided into low-risk and high-risk episodes.[14] We also analyze all non-reconciliation-eligible cancer types combined together. Depending on the model, this covariate is based on all 24 cancer episode types (along with the group of non-reconciliation-eligible cancers) or a subset of that are relevant to the outcome/subgroup.
	Previous episode	If beneficiaries with a current episode had an episode in the immediately preceding PP, they are flagged as having a previous episode.
	Chemotherapy source	Episodes are categorized based on the type(s) of chemotherapy the beneficiary used during the episode: Part B chemotherapy only, Part D chemotherapy only, or Part B and D chemotherapy.
	CMS HCC risk score	HCC score is used to quantify beneficiary severity of illness for their cancer and non-cancer comorbidities and predict plan payments in Medicare Advantage risk adjustment. HCC scores are based on beneficiary demographics and diagnostic history, including cancer and non-cancer codes. Each episode was assigned an HCC score based on the beneficiary's diagnosis information during the 12 months prior to the episode start date. For example, the HCC score for an episode that started on July 1, 2015 is constructed using diagnoses from July 1, 2014–June 30, 2015 claims. A beneficiary's HCC risk score for the episode is categorized based on quartiles. Quartile cut-points were derived from the episode-level distribution during the baseline period.
Practice-Level		
Practice organization and affiliations	Affiliation with an academic medical center	A practice is coded as affiliated if it was affiliated with an academic medical center. Affiliation is determined using Welch and Bindman 2016[15] and websites of medical school oncology/hematology departments, divisions, and institutes.

---

[13] Low- and high-intensity designations for prostate cancer follow the methodology used in the OCM PBP prediction model. Low-intensity prostate cancer episodes are defined as episodes in which the primary cancer type is prostate cancer and the patient is treated with androgen deprivation and/or an anti-androgen therapy, without any other chemotherapy during the episode. High-intensity prostate cancer episodes do not meet the above criteria.

[14] Low- and high-risk designations for bladder cancer episodes follow the methodology used in the OCM PBP prediction model. Specifically, low-risk bladder cancer episodes are defined as episodes in which the primary cancer type is bladder cancer and the patient is treated with intravesicular Bacillus Calmette-Guérin therapy and/or intravesicular mitomycin, without any other chemotherapy during the episode. High-risk bladder cancer episodes do not meet the above criteria.

[15] Welch, P. and Bindman, A.B. (2016). Town and gown differences among the largest medical groups in the US. Journal of Academic Medicine, July, 91(7):1007–14.

Domain	Model Covariate	Definition
	Affiliation with a health system	Practices are identified as affiliated with a health system based on information constructed from the July 2015, August 2016–2018 SK&A (now known as IQVIA) Office-Based Physician File for the baseline and intervention periods, respectively. The SK&A data are collected on a rolling basis via a telephone survey of physician practice sites. A practice is coded as affiliated if it was affiliated with at least one health system.
	Hospital ownership	Practices are identified as hospital-owned based on information constructed from the July 2015 and August 2016–2018 SK&A Office-Based Physician File for the baseline and intervention periods, respectively. The SK&A data are collected on a rolling basis via a telephone survey of physician practice sites. A practice is coded as owned if it was owned by at least one hospital.
Practice size and volume	Episode count	A practice's total number of episodes is categorized based on quartiles. Quartile cut-points are derived from the practice-level distribution during the baseline period.
	Practice size	Practices are coded as having 1–3 or 4+ oncology NPIs to distinguish between small and other practices.
Practice specialty type	Oncology-only specialty	Practices are coded as oncology-only if all NPIs within the practice had either an oncology specialty or a nurse practitioner/physician assistant (NP/PA) specialty.
	Presence of radiation oncology NPIs	A practice is flagged if it had at least one radiation oncology NPI.
	Presence of surgical oncology NPIs	A practice is flagged if it had a least one surgical oncology NPI.
	Presence of gynecologic oncology NPIs	A practice is flagged if it had a least one gynecologic oncology NPI.
	Percentage NP/PA NPIs	A practice's share of NPIs who is/are an NP/PA is categorized based on quartiles. Quartile cut-points were derived from the practice-level distribution during the baseline period.
Market-Level		
Market size	County population	The population size of the practice's county is categorized based on quartiles. For practices with multiple counties, this market characteristic and all others listed below are weighted according to the number of cancer E&M services the practice billed through each county. Quartile cut-points are derived from the market-level distribution during the baseline period.
Market demographics, income, and poverty	Percentage of population 65+	The percentage of population over age 65 in the practice's county is categorized based on quartiles. Quartile cut-points are derived from the market-level distribution during the baseline period.
	Percentage in poverty	The percentage of population living in poverty in the practice's county is categorized based on quartiles. Quartile cut-points are derived from the market-level distribution during the baseline period.
Market exposure to Medicare Alternative Models	Medicare Advantage penetration	The percentage of Medicare Advantage penetration in the practice's county is categorized based on quartiles. Quartile cut-points are derived from the market-level distribution during the baseline period.
Market provider supply	Percentage of population designated as a Primary Care HPSA	The practice's percentage of county population residing in a HPSA is categorized as 0 percent, >0–20 percent, or >20 percent. Cut-points are derived from the 2015 distribution of the HPSA proportion among markets with at least one OCM practice or comparison practice.

Domain	Model Covariate	Definition
	Ratio of specialists to primary care providers	A ratio is calculated from the number of specialists divided by the number of primary care physicians in the practice's county. Each practice's ratio is categorized based on quartiles. Quartile cut-points are derived from the market-level distribution during the baseline period.
Market health services utilization	Total emergency department (ED) visits among FFS population	The practice's county-level IP emergency department visits per 10,000 FFS population is categorized based on quartiles. Quartile cut-points are derived from the market-level distribution during the baseline period (composite score averaging 2014 and 2015 values).

Subgroup Analyses

We conduct subgroup analyses for a select group of outcome measures to examine differential impacts of OCM by cancer episode type. The subgroup analyses serve several purposes: (1) to inform the generalizability of OCM, (2) to identify underlying drivers of success in OCM, and (3) to measure whether OCM leads to unintended consequences for particular groups of beneficiaries.

We identified two subgroup categories: cancer treatment intensity (i.e., higher-risk and lower-risk episodes) and individual cancer episode type. The specific subgroups are shown in **Exhibit A-6** below. We ran DID analyses for the specific subgroup samples and compare results across each subgroup category. Outcome measures for which we conducted subgroup analyses included: TEP, Part A payments, Part B payments, Part D payments, Part B chemotherapy payments, Part B non-chemotherapy drug payments, Part B imaging payments, and ACH payments. We did not run DID analyses for every outcome measure and subgroup combination.

Exhibit A-6: Subgroups Evaluated in the Report Covering PP1–6

Subgroup Category	Episode Subgroups
Cancer episode type	Low-risk breast cancer Low-intensity prostate cancer High-risk breast cancer Lung cancer Lymphoma Colorectal/small intestine cancer Multiple myeloma Non-reconciliation-eligible cancers High-intensity prostate cancer Chronic leukemia
Episode risk group (i.e., treatment Intensity)	Lower-risk episodes[16] Higher-risk episodes[17]

Parallel Trends Assumption

DID analysis assumes that trends for outcome measures in the baseline period were similar for OCM and comparison episodes, and would have remained so in the absence of OCM. Thus, DID accounts for unobserved variables affecting both groups equally, which are assumed to remain equally relevant for both groups over time. Failure of the baseline (pre-OCM) parallel trends assumption results in biased DID estimates.

For each outcome measure, we test the null hypothesis that episodes attributed to OCM practices and comparison practices had parallel trends during the baseline period. We compare baseline trends on a quarterly basis instead of a PP basis. For each measure, we estimate a DID regression model using the same functional form and covariates as the main impact analyses, including an indicator for OCM versus comparison, a linear trend, and an OCM-specific trend. We reject the null hypothesis that there were parallel trends in the baseline (i.e., cannot conclude that trends were parallel) at the 5 percent level of significance. For outcome measures assessed for a subpopulation of the data (e.g., cancer type), we limit the episode sample to the subgroup of interest, and ran an analogous parallel trends test.

Among outcome measures for which we reject the null hypothesis, we further review the data to determine whether OCM and comparison baseline trends appear visually parallel, and whether the removal of a handful of extreme values would result in the outcome measure passing the parallel trends test (i.e., we cannot reject the null hypothesis). Using this combination of criteria, we identified the set of outcome measures (and relevant subgroups, where applicable) that we deem cannot be reliably reported due to a potential bias in the DID estimate. This report does not include results for these outcome measures and subgroup combinations.

Sensitivity Tests

We perform several sensitivity tests to understand whether the reported impact estimates were robust with respect to the model specification and the episode sample used, and perform this testing on four outcome measures: TEP, Part A payments, Part B payments without MEOS, and Part D payments. These measures were selected because they are important for understanding the impact of OCM, because they rely on

---

[16] Lower-risk cancer episodes include low-risk breast cancer, low-intensity prostate cancer and low-risk bladder cancer.

[17] Higher-risk cancer episodes include the 21 cancer types and non-reconciliation-eligible cancers not included in the lower-risk cancer type subgroup.

different types of data and have different functional forms. We conduct sensitivity tests for the full sample of episodes and also for the subsamples of higher-risk and lower-risk episodes, separately.

The tests examine sensitivity of the results to the following:

- Choice of model functional form
- Exclusion of episodes with extreme large payment values (top 5 and 10 percent of TEP)
- Exclusion of episodes for the two largest OCM practices and practices that are part of the U.S. Oncology Network
- Exclusion of episodes for beneficiaries without Part D enrollment in all months
- Exclusion of episodes for specific cancer episode types, or with specific treatment timing (e.g., new versus ongoing chemotherapy or hormonal therapy treatment)

Estimation of Net Impact to Medicare

A reduction in per-episode payments (TEP) implies that OCM is reducing episode-level spending, but this does not necessarily translate into net savings for Medicare because TEP does not include the MEOS payment or PBP that Medicare pays to participating practices. To assess the net impact of OCM, we must include the MEOS payments and PBP made to participating practices to determine whether OCM is achieving sufficient savings to cover its costs. To calculate the net impact to Medicare in PP1 to PP5, we add total MEOS and PBP amounts paid by Medicare to the gross reduction in episode payments measured by TEP, as follows:

$$Net\ Impact = (Gross\ Impact\ on\ TEP) + (MEOS + PBP)$$

Using our DID estimates for TEP in each PP, we multiply TEP by the number of OCM episodes in that PP to estimate the gross impact on TEP. We then sum MEOS payments and PBP with the gross impact on TEP, to estimate the net impact for Medicare.

For PP3–PP5, we also calculate the impact on Medicare spending separately among lower-risk and higher-risk episodes. Since PBP is paid to practices and not defined for each episode, we only include MEOS payments and do not include PBP in the savings/losses estimates for higher-risk and lower-risk episodes. **Exhibit A-7** defines the measures in this analysis.

Exhibit A-7: Definition of Measures Used in the Estimation of the Net Impact to Medicare

Measure	Description
Episode-level DID estimate of TEP, by PP	A per episode estimate of the impact on TEP attributable to the OCM model. Estimated for each PP.
Total number of episodes attributed to OCM participants, by PP	The number of program episodes attributed to OCM participants for each PP separately. This count includes reconciliation- and non-reconciliation-eligible episodes.
Gross impact on TEP, by PP	The product of the DID estimate of TEP by the total number of episodes, calculated for each PP separately.
MEOS + PBP, by PP	Sum of MEOS and PBP paid amounts for each PP separately (first true-up reconciliation results).
Net impact to Medicare, by PP	Gross impact on TEP + total MEOS + PBP, calculated for each PP separately.

Notes: DID: difference-in-difference, TEP: Total episode payments, PP: performance period, MEOS: monthly enhanced oncology services, PBP: performance-based payment.

B. Payment and Utilization Outcome Analyses

B.1. Impact on TEP: Overall and by Episode Risk Group and Cancer Type

Exhibit B-1: OCM Reduced TEP Overall and for Higher-Risk Episodes, but Increased TEP for Lower-Risk Episodes

TEP Reduction for Higher-Risk Episodes Was Concentrated in High-Risk Breast, Lung, Lymphoma, and Colorectal Cancer Episodes

TEP	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
All Episodes	$28,760	$34,048	$28,482	$34,069	-$298**	-1.0%	-$86	-$297**	-$332**	-$375**	-$389*	-$309
Episode Risk Group												
Lower-risk episodes	$7,239	$7,597	$7,337	$7,564	$130*	1.8%	$82	$215*	$213**	$161*	$60	$50
Higher-risk episodes	$40,024	$47,875	$39,504	$47,843	-$487**	-1.2%	-$141	-$523**	-$594***	-$632**	-$582**	-$452
Cancer Type												
Low-risk breast cancer	$5,376	$5,562	$5,453	$5,613	$27	0.5%	$14	$116	$158*	$100	-$23	-$202**
Low-intensity prostate cancer	$11,352	$12,046	$11,314	$11,771	$237	2.1%	$120	$394	$181	$282	$100	$325
High-risk breast cancer	$35,631	$41,710	$34,526	$41,490	-$885***	-2.5%	-$746**	-$606	-$564	-$1,131***	-$798*	-$1,471***
Lung cancer	$39,934	$53,197	$39,270	$53,644	-$1,112***	-2.8%	-$621	-$1,053**	-$1,204***	-$1,500***	-$1,954***	-$394
Lymphoma	$43,634	$49,214	$44,249	$50,763	-$934*	-2.1%	-$585	-$454	-$1,202*	-$1,271	-$1,526*	-$641
Colorectal/small intestine cancer	$36,021	$36,330	$35,054	$36,228	-$865*	-2.4%	-$401	-$565	-$1,528***	-$925	-$1,179*	-$678
Multiple myeloma	$53,713	$71,961	$53,558	$71,843	-$36	-0.1%	$567	-$436	-$43	$432	-$10	-$692
Non-reconciliation-eligible cancers	$37,600	$46,531	$35,836	$44,496	$271	0.7%	-$258	$742	-$135	$222	$178	$758
High-intensity prostate cancer	$42,178	$46,784	$41,936	$46,917	-$376	-0.9%	$250	-$274	-$320	-$1,064	-$490	-$299
Chronic leukemia	$44,217	$49,520	$43,996	$48,817	$483	1.1%	-$176	-$638	$58	$879	$1,833**	$1,446

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01 **Source**: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

B.2. Net Impact of OCM

Exhibit B-2: OCM Resulted in Net Losses to Medicare in Every PP
Losses were Lowest in PP5 at $61M

PP	Total Program Episodes	Gross Impact on TEP			+ PBP Payments	+ MEOS Payments	= Net Impact (Losses to Medicare)		
		Estimate	LCL	UCL			Estimate	LCL	UCL
PP1	139,667	-$11,956,182	-$42,126,697	$18,214,334	$14,295,955	$98,575,061	$100,914,834	$70,744,320	$131,085,350
PP2	132,629	-$39,454,598**	-$72,325,563	-$6,583,631	$17,708,460	$93,880,339	$72,134,201**	$39,263,236	$105,005,168
PP3	128,724	-$42,750,761**	-$74,611,507	-$10,890,018	$19,031,892	$89,464,798	$65,745,929**	$33,885,183	$97,606,672
PP4	133,202	-$49,992,973**	-$89,453,037	-$10,532,904	$33,297,129	$94,134,524	$77,438,679**	$37,978,615	$116,898,748
PP5	129,098	-$50,183,057*	-$92,889,407	-$7,476,708	$22,160,233	$88,893,894	$60,871,070*	$18,164,720	$103,577,419

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019; OCM first true-up reconciliation reports, PP1–PP5.
Notes: LCL: lower confidence limit. UCL: upper confidence limit.

B.3. Impact on Payments by Medicare Coverage Part: Overall and by Episode Risk Group and Cancer Type

Exhibit B-3: OCM's Reduction in TEP Was Driven by Part A and Part B Payments

Measure	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
TEP without MEOS	$28,760	$34,048	$28,482	$34,069	-$298**	-1.0%	-$86	-$297**	-$332**	-$375**	-$389*	-$309
Part A payments	$6,070	$5,924	$5,946	$5,905	-$104*	-1.7%	-$68	-$133*	-$160**	-$131*	-$78	-$61
Part B payments	$17,096	$20,381	$16,928	$20,396	-$182*	-1.1%	-$61	-$175	-$158	-$277**	-$210	-$208
Part D payments[a]	$6,679	$9,323	$6,731	$9,336	$39	0.6%	$74	$61	$25	$117	-$77	$32

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: TEP: Total episode payments. [a] Part D payments are calculated as the sum of low-income cost-sharing and reinsurance amounts, as reflected on the PDE. MEOS: Monthly Enhanced Oncology Services payment. OCM: OCM intervention group; COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-4: OCM Had No Overall Impact on Part D Payments or Part D Gross Drug Costs

Measure	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Part D payments[a]	$6,679	$9,323	$6,731	$9,336	$39	0.6%	$74	$61	$25	$117	-$77	$32
Part D GDC[b]	$10,386	$13,850	$10,475	$13,901	$38	0.4%	$78	$115	$22	$151	-$129	-$11

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: [a] Part D payments are calculated as the sum of low income cost-sharing and reinsurance amounts paid by Medicare, as reflected on the PDE. [b] Part D gross drugs costs (GDC) is calculated as the sum of ingredient cost, dispensing fee, vaccine administration fee, and sales tax, as shown on the PDE, reflecting the total spending on the prescription fill from all parties. OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-5: OCM Reduced Part A Payments for Higher-Risk Episodes, but Had No Impact on Episodes for Most of the Common Cancers

Part A Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$2,294	$2,222	$2,247	$2,144	$32	1.4%	$18	$63	$150**	$18	$16	-$69
Higher-risk episodes	$8,018	$7,854	$7,848	$7,869	-$185**	-2.3%	-$105	-$226**	-$344***	-$235**	-$148	-$66
Cancer Type												
Low-risk breast cancer	$1,680	$1,638	$1,666	$1,621	$3	0.2%	-$17	$98	$120**	$50	-$81	-$153**
Low-intensity prostate cancer	$3,604	$3,430	$3,460	$3,195	$91	2.5%	$53	$26	$208	$18	$239	$22
High-risk breast cancer	$4,986	$4,769	$4,938	$4,656	$66	1.3%	$145	-$20	$220	-$9	$94	-$31
Lung cancer	$9,410	$9,078	$9,119	$8,990	-$204	-2.2%	-$109	-$57	-$399*	-$260	-$453*	$29
Lymphoma	$7,633	$7,652	$7,522	$7,908	-$367	-4.8%	-$358	-$194	-$781**	-$385	-$264	-$231
Multiple myeloma	$7,675	$7,073	$7,861	$7,547	-$288	-3.8%	-$151	-$370	-$495	-$118	-$344	-$260
Non-reconciliation-eligible cancers	$7,441	$7,247	$7,159	$7,206	-$241	-3.2%	-$524	$112	-$761**	-$101	-$361	$97
High-intensity prostate cancer	$6,376	$5,888	$5,974	$5,661	-$174	-2.7%	-$26	-$69	-$130	-$876***	-$153	$224
Chronic leukemia	$5,161	$5,047	$5,278	$4,953	$211	4.1%	-$172	$262	$238	$452	$650	-$92

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01
Source: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group; COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences. Colorectal cancer/small intestine is not included in this table because the Part A Payments impact estimate could not be reliably reported due to failure of the parallel trends assumption.

Exhibit B-6: OCM Reduced Part B Payments for Higher-Risk Episodes, but Slightly Increased Part B Payments for Lower-Risk Episodes

OCM Reduced Part B Payments for High-Risk Breast, Lung, Colorectal, and High-Intensity Prostate Cancer Episodes

Part B Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$4,474	$4,860	$4,597	$4,901	$81*	1.8%	$37	$123	$49	$134**	$41	$98*
Higher-risk episodes	$23,550	$28,460	$23,336	$28,540	-$294**	-1.2%	-$84	-$327*	-$271	-$454**	-$297	-$324
Cancer Type												
Low-risk breast cancer	$3,144	$3,323	$3,204	$3,388	-$6	-0.2%	-$10	-$14	$17	$15	$33	-$73
Low-intensity prostate cancer	$7,438	$8,258	$7,565	$8,232	$153	2.1%	$60	$333	-$25	$299*	-$78	$293**
High-risk breast cancer	$24,886	$27,610	$24,221	$27,807	-$861***	-3.5%	-$833***	-$623**	-$654**	-$1,001***	-$669*	-$1,388***
Lung cancer	$27,166	$39,215	$26,787	$39,534	-$697**	-2.6%	-$273	-$713	-$711	-$1,106**	-$1,306***	-$117
Lymphoma	$30,958	$35,821	$31,606	$36,709	-$240	-0.8%	$119	-$118	-$262	-$941*	-$514	$237
Colorectal/small intestine cancer	$25,956	$25,967	$25,171	$26,049	-$867**	-3.3%	-$335	-$683	-$1,140**	-$811	-$1,100**	-$1,196**
Multiple myeloma	$22,050	$28,000	$21,697	$27,777	-$130	-0.6%	$70	-$322	$108	-$10	-$32	-$561
High-intensity prostate cancer	$18,108	$19,043	$17,620	$19,288	-$733*	-4.0%	-$326	-$1,292**	-$591	-$853	-$605	-$749
Chronic leukemia	$12,792	$14,107	$12,739	$13,997	$58	0.5%	-$112	-$285	$154	$49	$661**	-$9

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences. Non-reconciliation-eligible cancers are not included in this table because the Part B payments impact estimate could not be reliably reported due to failure of the baseline parallel trends assumption.

Exhibit B-7: OCM Increased Part D Payments for High-Intensity Prostate Cancer Episodes, but Had No Impact on Other Episode Types

Part D Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$528	$578	$550	$581	$19	3.6%	$32	$33	$14	$11	-$1	$23
Higher-risk episodes	$10,526	$14,514	$10,406	$14,328	$66	0.6%	$83	$96	$90	$188	-$105	$42
Cancer Type												
Low-risk breast cancer	$552	$601	$585	$605	$30	5.4%	$40	$33	$22	$35	$23	$24
Low-intensity prostate cancer	$464	$534	$429	$514	-$15	-3.1%	$11	$46	-$14	-$53	-$98*	$13
High-risk breast cancer	$7,048	$11,614	$6,639	$11,212	-$8	-0.1%	-$14	$127	-$69	$4	-$169	$65
Lung cancer	$4,375	$6,464	$4,419	$6,732	-$223	-5.1%	-$283	-$316	-$96	-$108	-$220	-$304
Lymphoma	$6,662	$7,638	$6,799	$8,135	-$360	-5.4%	-$409	-$117	-$141	$101	-$903**	-$707
Colorectal/small intestine cancer	$2,591	$2,986	$2,509	$2,766	$138	5.3%	$150	$477**	$61	$229	-$9	-$111
Multiple myeloma	$27,836	$43,076	$27,926	$42,559	$607	2.2%	$779*	$367	$521	$925	$566	$484
Non-reconciliation-eligible cancers	$15,383	$20,088	$14,300	$18,448	$556	3.6%	$325	$632	$719	$949**	$418	$329
High-intensity prostate cancer	$19,788	$24,522	$20,518	$24,622	$630*	3.2%	$800**	$1,184***	$428	$810	$384	$264
Chronic leukemia	$28,525	$33,003	$28,188	$32,423	$243	0.9%	$116	-$641	-$354	$410	$599	$1,649**

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01. OCM: OCM intervention group.
Source: Medicare claims 2014–2019.
Notes: Part D payments are calculated as the sum of low-income cost-sharing and reinsurance amounts, as reflected on the PDE. OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

B.4. Impact on Part A and Part B Payment Components

Exhibit B-8: OCM Had No Overall Impact on Key Part A Payment Components

Measure	% of Part A Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
		Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Part A Payments	**100.0%**	**$6,070**	**$5,924**	**$5,946**	**$5,905**	**-$104***	**-1.7%**	**-$68**	**-$133***	**-$160****	**-$131***	**-$78**	**-$61**
ACH payments	65.3%	$3,961	$3,921	$3,707	$3,630	$37	0.9%	$46	-$1	-$20	$41	$83	$69
SNF payments	11.0%	$669	$623	$633	$596	-$9	-1.3%	$8	-$13	-$22	-$17	-$25	$14
HHA payments	10.9%	$663	$613	$654	$615	-$12	-1.8%	-$16	-$1	-$16	-$22	-$1	-$16
Hospice payments	7.7%	$468	$450	$427	$412	-$3	-0.7%	$7	-$1	$11	-$17	-$6	-$13
IRF payments	3.5%	$215	$235	$190	$207	$3	1.5%	-$3	$4	$1	-$8	$3	$21*
LTCH payments	2.0%	$121	$89	$117	$84	$2	1.3%	$9	$9	$7	$6	-$12	-$9

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: ACH: acute care hospital. OIP: other inpatient facility. SNF: skilled nursing facility. HHA: home health agency. IRF: inpatient rehabilitation facility. LTCH: long-term care hospital. OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-9: OCM Reduced Part B Payments, Primarily for Non-Chemotherapy Drugs

Measure	% of Part B Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
		Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Part B Payments	100.0%	$17,096	$20,381	$16,928	$20,396	-$182*	-1.1%	-$61	-$175	-$158	-$277**	-$210	-$208
Chemo payments	44.9%	$7,674	$10,660	$7,547	$10,524	$9	0.1%	$55	-$62	$51	-$83	$2	$90
Other payments without MEOS	15.9%	$2,721	$2,787	$2,841	$2,934	-$27	-1.0%	-$16	-$26	-$36	-$23	-$4	-$53
Non-chemo drug payments	15.6%	$2,672	$2,815	$2,450	$2,754	-$161***	-6.0%	-$90*	-$114**	-$158***	-$156**	-$209***	-$239***
Non-cancer E&M payments	5.3%	$898	$904	$879	$892	-$6	-0.7%	-$10	-$1	-$8	-$13	-$3	-$2
Radiation therapy payments	4.8%	$817	$816	$912	$899	$13	1.6%	-$1	$21	$13	$20	$15	$7
Imaging payments	4.8%	$814	$834	$814	$853	-$19***	-2.3%	-$10	-$10	-$19**	-$27***	-$23***	-$25***
Chemo administration payments	3.7%	$629	$667	$666	$695	$8	1.3%	$6	$12	$9	$7	$9	$7
Labs payments	2.6%	$451	$486	$414	$449	$0	0.1%	$4	$6	-$2	-$6	$0	$1
Cancer E&M payments	2.3%	$389	$377	$353	$338	$3	0.8%	-$0	$5	$2	$7	$4	$1

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: E&M: evaluation and management. MEOS=Medicare Enhanced Oncology Service payment. OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-10: OCM Had No Impact on ACH Payments for Higher-Risk or Lower-Risk Episodes

OCM Increased ACH Payments for High-Risk Breast Cancer Episodes but Had No Impact for Other Common Episode Types

Part A ACH Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$1,289	$1,286	$1,248	$1,204	$41	3.1%	$39	$26	$101**	$52	$32	-$3
Higher-risk episodes	$5,336	$5,288	$4,973	$4,902	$24	0.5%	$49	-$15	-$103	$14	$94	$99
Cancer Type												
Low-risk breast cancer	$922	$938	$927	$908	$36	3.9%	$9	$46	$90**	$76*	$20	-$23
Low-intensity prostate cancer	$2,026	$1,950	$1,884	$1,779	$29	1.4%	$44	-$4	$102	$22	$47	-$27
High-risk breast cancer	$2,995	$2,903	$2,926	$2,695	$139*	4.7%	$156	$67	$245**	$129	$224	$18
Lung cancer	$6,071	$5,908	$5,745	$5,623	-$42	-0.7%	-$59	$13	-$157	$45	-$111	$11
Lymphoma	$5,596	$5,681	$5,158	$5,339	-$96	-1.7%	-$183	$106	-$625**	-$70	-$22	$215
Colorectal/small intestine cancer	$5,344	$5,434	$4,958	$5,030	$18	0.3%	-$20	-$125	-$309	-$45	$112	$499**
Multiple myeloma	$5,432	$5,015	$5,329	$4,943	-$31	-0.6%	$83	-$124	-$259	$149	$5	-$51
Non-reconciliation-eligible cancers	$4,831	$4,832	$4,375	$4,202	$174	3.6%	$162	$453*	-$285	$190	$80	$395
High-intensity prostate cancer	$3,663	$3,499	$3,436	$3,210	$61	1.7%	$100	$151	$164	-$389*	$219	$146
Chronic leukemia	$3,536	$3,467	$3,411	$3,261	$81	2.3%	-$130	$300	$191	$248	$12	-$169

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: ACH: acute care hospital. OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-11: OCM Had No Impact on Part B Chemotherapy Payments for Lower-Risk or Higher-Risk Episodes

High-Risk Breast Cancer Episodes Were the Only Common Episode Type for Which OCM Had an Impact, Reducing Payments by $500 per Episode

Part B Chemotherapy Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$351	$352	$351	$354	-$1	-0.3%	$4	$0	$0	-$2	-$4	-$5
Higher-risk episodes	$11,439	$16,035	$11,307	$15,856	$47	0.4%	$105	-$80	$88	-$74	$55	$182
Cancer Type												
Low-intensity prostate cancer	$1,146	$1,145	$1,148	$1,153	-$6	-0.5%	$10	-$1	-$6	-$10	-$12	-$16
High-risk breast cancer	$13,058	$15,156	$12,386	$14,986	-$502**	-3.8%	-$525**	-$300	-$449*	-$708**	-$311	-$712**
Lung cancer	$12,756	$25,335	$12,484	$25,255	-$192	-1.5%	-$116	-$487	$4	-$708	-$735	$822
Lymphoma	$19,772	$22,837	$20,139	$23,567	-$363	-1.8%	-$59	-$233	-$329	-$1,008**	-$539	-$45
Colorectal/small intestine cancer	$11,985	$12,253	$11,863	$12,085	$46	0.4%	-$93	-$294	-$35	$294	$230	$199
Multiple myeloma	$13,151	$18,879	$12,856	$18,648	-$64	-0.5%	$80	-$391	$194	$63	$65	-$385
Non-reconciliation-eligible cancers	$6,746	$11,648	$6,454	$11,192	$164	2.4%	$241	$297	$153	-$234	$80	$447
High-intensity prostate cancer	$6,408	$6,851	$6,228	$6,950	-$279	-4.3%	-$288	-$842	-$342	$22	-$252	-$53
Chronic leukemia	$6,531	$7,135	$6,274	$7,031	-$153	-2.3%	-$396*	-$499*	-$100	-$5	$217	-$5

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-12: OCM Reduced Part B Non-Chemotherapy Drug Payments for Higher-Risk Episodes

Reductions Were Concentrated in High-Risk Breast, Lung, Colorectal, and High-Intensity Prostate Cancer Episodes

Part B Non-Chemo Drug Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$621	$735	$547	$642	$19	3.1%	$5	$80	-$13	$52	-$27	$15
Higher-risk episodes	$3,683	$3,883	$3,433	$3,888	-$256***	-6.9%	-$136**	-$220***	-$237***	-$267***	-$300***	-$366***
Cancer Type												
Low-risk breast cancer	$322	$379	$327	$377	$7	2.1%	$9	$2	$30	$24	-$5	-$19
Low-intensity prostate cancer	$1,340	$1,545	$1,149	$1,298	$57	4.3%	-$42	$243	-$98	$135	-$43	$124
High-risk breast cancer	$4,292	$4,724	$4,143	$4,867	-$292***	-6.8%	-$216**	-$251**	-$226**	-$190	-$353**	-$528***
Lung cancer	$4,238	$3,735	$3,829	$3,626	-$300**	-7.1%	-$108	-$226*	-$312**	-$199	-$397**	-$542***
Lymphoma	$4,367	$5,920	$4,560	$5,887	$227	5.2%	$134	$172	$100	$224	$219	$523
Colorectal/small intestine cancer	$4,622	$4,269	$4,011	$4,214	-$556**	-12.0%	$7	-$190	-$559*	-$745**	-$859***	-$1,055***
Multiple myeloma	$1,942	$2,223	$1,709	$2,078	-$89	-4.6%	-$52	-$41	-$116	-$136	-$119	-$67
Non-reconciliation-eligible cancers	$3,063	$3,126	$2,869	$3,158	-$226*	-7.4%	-$307**	-$330*	-$147	-$291*	-$52	-$253
High-intensity prostate cancer	$5,736	$6,043	$5,278	$6,007	-$422**	-7.4%	-$66	-$351	-$162	-$700***	-$488*	-$697***
Chronic leukemia	$1,589	$2,233	$1,691	$2,230	$105	6.6%	$167	$48	$54	$89	$234	$53

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

Exhibit B-13: OCM Led to a Small Decrease in Part B Imaging Payments for Higher-Risk Episodes

Part B Imaging Payments	OCM		COMP		Impact Estimates Through PP6		Period by Period Impact Estimates					
	Baseline Mean	Int. Mean	Baseline Mean	Int. Mean	DID	Percentage Change	PP1 DID	PP2 DID	PP3 DID	PP4 DID	PP5 DID	PP6 DID
Episode Risk Group												
Lower-risk episodes	$378	$390	$380	$394	-$2	-0.6%	$1	$0	$3	-$5	-$1	-$10
Higher-risk episodes	$1,040	$1,066	$1,039	$1,092	-$26***	-2.5%	-$15	-$14	-$30***	-$37***	-$32***	-$29**

Asterisks denote statistically significant impact estimates at *p<0.10, **p<0.05, and ***p<0.01.
Source: Medicare claims 2014–2019.
Notes: OCM: OCM intervention group. COMP: comparison group. Int.: intervention period. PP: performance period. DID: difference-in-differences.

B.5.   Results of Sensitivity Analyses

As discussed in **Appendix A**, we performed sensitivity tests on key payment outcome measures to assess whether impact estimates were robust to changes in model specification and/or the types of practices and episodes included in the sample. We conducted this analysis for all episodes, and separately for higher-risk and lower-risk cancer episodes. A detailed list of tests is in **Exhibit B-14**.

Exhibit B-14: Sensitivity Tests Conducted for Selected Payment Outcome Measures

Sensitivity Test	Outcome Measures			
	TEP	Part A Pmts	Part B Pmts	Part D Pmts
Model Specification				
Alternative model functional form—two-part model (TPM) instead of ordinary least squares (OLS)		X		X
Sample Criteria				
Exclude positive outliers: episodes with payments 10SD above the mean of the i) overall distribution and ii) cancer-specific distribution	X	X	X	X
Exclude positive outliers: of episodes with payments 5 standard deviations above the mean of the i) overall distribution and ii) cancer-specific distribution	X	X	X	X
Exclude positive outliers: of episodes with payments in the top 1% of the distribution of each cancer type	X	X	X	X
Exclude the two largest OCM practices	X	X	X	X
Exclude episodes for beneficiaries not enrolled in Part D for all months of the episode	X	X	X	
Include episodes for beneficiaries not enrolled in Part D for all months of the episode				X
Exclude episodes for which the beneficiary had ongoing chemotherapy (i.e., an episode in the previous PP)	X	X	X	X
Exclude episodes with inpatient or outpatient CAR-T cell therapy	X	X	X	

We expected some variability in the magnitude and precision of the impact estimates due to changes in specification and sample composition. Therefore, in assessing the results of the sensitivity tests, we used a two-stage approach: we first determined whether the sensitivity impact estimates were meaningfully different from the main impact estimates. If they were, we evaluated whether such differences affected the reliability or interpretation of the main estimates.

To determine potentially meaningful differences, we assessed whether the difference between the main impact estimate and the impact estimate obtained from a given sensitivity test was statistically different from zero. In addition to the traditional statistical threshold of $p<0.1$, we evaluated several alternative thresholds (including $p<0.2$, $p<0.3$, $p<0.4$, and $p<0.5$), since $p<0.1$ may be too stringent to identify meaningful differences in sensitivity estimates. Using this strategy, we erred on the side of identifying a larger number of potentially meaningful differences that we could subsequently review to evaluate the reliability of the main impact estimate.

Sensitivity Results: All Episodes

For the full episode sample, the impact estimates of the key payment outcomes were generally consistent across different model specifications and sample criteria. Across all outcomes, the test results were not systematically different in both sign and statistical significance from the main impact estimates. The vast majority of statistical tests produced impact estimates that were also not meaningfully different from their

corresponding main impact estimates (i.e., difference between main and sensitivity estimate was not statistically significantly different from zero at values of alpha less than 0.4.).

For Part D payments, the sensitivity tests systematically produced small (i.e, >0) impact estimates that were similar to the main Part D payment impact estimate. The exception was the alternative functional form test, which used a TPM instead of OLS. The impact estimate obtained from TPM was positive ($121) and statistically significant (p<0.10). In contrast, the main Part D estimate from OLS was smaller ($39) and not statistically significant. Despite this difference in magnitude between the OLS and TPM estimates, the difference was statistically significant only at p<0.5. Additionally, we chose OLS as the functional form for the main impact estimates due to certain advantages of using OLS. In particular, OLS allows easier interpretation and comparison of estimates, and makes it easier to "add up" component impact estimates (e.g., Part A, Part B, Part D payments) into the overall estimate (e.g., TEP). Additionally, the use of OLS enables consistency within subgroup categories (i.e., TPM will not converge when we estimate Part D impact for each of the common cancer episode types).

Sensitivity Results: Higher- and Lower-Risk Episodes

We also performed the sensitivity tests for the subgroups of higher-risk and lower-risk episodes. We identified meaningful differences between the sensitivity tests and the main estimates in a few cases. **Exhibit B-15** shows which sensitivity tests produced potentially meaningful changes in results as compared with the main estimate, by payment outcome and subgroup.

Exhibit B-15: Sensitivity Tests That Yielded Meaningful Differences from the Main Impact Estimate, by Payment Outcome Measure, for Higher-Risk and Lower-Risk Episodes

Outcome	Higher-Risk Episodes	Lower-Risk Episodes
TEP	No meaningful differences from main estimate	Sensitive to exclusion of episodes with ongoing treatment
Part A payments	No meaningful differences from main estimate	No meaningful differences from main estimate
Part B payments	No meaningful differences from main estimate	Sensitive to exclusion of episodes with ongoing treatment
Part D payments	Sensitive to model functional form (TPM)	No meaningful differences from main estimate

For higher-risk episodes, the estimated impact on Part D payments was sensitive to the use of TPM instead of OLS, as was true for the full sample. As discussed, we intentionally chose OLS due to its relative advantages over TPM. We are confident in the reliability and interpretability of the main payment impact estimates for higher-risk episodes.

For lower-risk episodes, the impact estimates for both TEP and Part B payments were sensitive to the exclusion of episodes with ongoing chemotherapy (i.e., episodes for beneficiaries who had had an episode in the immediately preceding PP). Episodes with ongoing treatment have different costs and clinical severity than newly initiated episodes; as a result, we excluded episodes with ongoing treatment to understand how much they were driving the main findings. About half of the lower-risk episodes were for beneficiaries who had had an episode in the previous PP. As a result, the already (relatively) small subsample of lower-risk episodes was further reduced in size after excluding these episodes for ongoing chemotherapy. Both the impact estimates and the corresponding standard errors for that reduced sample were larger in absolute value compared with the main estimates. However, the sign and statistical significance of the sensitivity estimates were similar to the main lower-risk impact estimate. Additionally, the sensitivity estimates were only marginally statistically different from the main estimates (i.e., different only at p<0.5). As a result, we deem the test results to be reasonably consistent with the main results reported in this annual report.