Evaluation of the Maryland Total Cost of Care Model: Progress Report Appendices

April 2024

Greg Peterson, Jason Rotter, Rachel Machta, Keri Calkins, Katie Morrison Lee, Amanda Markovitz, Rumin Sarwar, Kate Stewart, Jake Vogler, Isabel Platt, Tyler Rose, Danielle Whicher, Nancy McCall

Submitted to:

U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services
7500 Security Blvd.
Baltimore, MD 21244-1850
Contracting Officer's Representative: Caroline Ly
Contract Number: HHSM-500-2014-00034I

Submitted by:

Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221
Phone: (202) 484-9220

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2014-00034I, Task Order 75FCMC19F0003. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. Mathematica assumes responsibility for the accuracy and completeness of the information contained in this report.

Contents

Acronyms

ACO	Accountable Care Organization
ADI	Area Deprivation Index
CMF	care management fee
CTI	Care Transformation Initiative
CTO	Care Transformation Organization
CMS	Centers for Medicare & Medicaid Services
CRISP	Chesapeake Regional Information System for our Patients
CHIP	Children’s Health Insurance Program
CMM	comprehensive medication management
CPCP	Comprehensive Primary Care Payment
CI	confidence interval
DRG	diagnosis-related group
EHR	electronic health record
ED	emergency department
ESRD	end-stage renal disease
ECIP	Episode Care Improvement Program
EQIP	Episode Quality Improvement Program
FQHC	Federally Qualified Health Center
FFS	fee for service
HEART	Health Equity Advancement Resource and Transformation
HSCRC	Health Services Cost Review Commission
HCPCS	Healthcare Common Procedure Coding System
HRSN	health-related social need
HCC	Hierarchical Condition Category
ICC	intracluster correlation coefficient
IPPS	Inpatient Prospective Payment System
MDAPM	Maryland All-Payer Model

MDPCP	Maryland Primary Care Program
PUMA	Maryland Public Use Microdata Area
MD TCOC	Maryland Total Cost of Care Model
MAX	Medicaid Analytic eXtract
MD-PPAS	Medicare Data on Provider and Specialty
MPA	Medicare Performance Adjustment
NPPES	National Plan and Provider Enumeration System
NPI	National Provider Identifier
PFACs	Patient and Family Advisor Councils
PQI	Prevention Quality Indicator
PCPs	primary care providers
PPS	prospective payment system
PHE	public health emergency
SSP	Shared Savings Program
SVI	Social Vulnerability Index
SIHIS	Statewide Integrated Health Improvement Strategy
TIN	Tax Identification Number
TAF	T-MSIS Analytic File
T-MSIS	Transformed Medicaid Statistical Information System

Appendix A. Methods for Calculating Participation in the Components of the Maryland Total Cost of Care Model and Their Incentives

In Chapter 1 of the Progress Report, we described each component of the Maryland Total Cost of Care (MD TCOC) Model on three dimensions: number of participants, reach, and the size of the incentives and supports in 2022 (see Exhibits 1.2 and 1.3). This appendix details how we calculated these descriptive statistics. As in Chapter 1, we group the components by the size of the incentives or investments in 2022 (the numbering of the components in this appendix matches the numbering in Exhibits 1.2 and 1.3 in Chapter 1 of the main report). We also describe how the outcomes we have included when estimating impacts of the model align with cost and quality goals described in the (a) state agreement establishing the MD TCOC Model, and (b) the Statewide Integrated Health Improvement Strategy (SIHIS).

A.1. Components with the largest incentives or investment under the MD TCOC Model in 2022 (more than $100 million)

A.1.1. Component 1: All-payer hospital global budgets

A.1.1a. Participation

The agreement between Maryland and the Centers for Medicare & Medicaid Services (CMS) establishing the MD TCOC Model lists the 52 participating hospitals (HSCRC 2018) (see Exhibit A.1). These are the hospitals for which Health Services Cost Review Commission (HSCRC) sets the rates that all payers, including Medicare, pay for hospital care in the state. These 52 hospitals include 44 acute care hospitals, seven freestanding emergency centers or freestanding medical facilities (some of which also offer observation services and outpatient surgeries), and one specialty emergency department (University of Maryland Shock Trauma Center). The MD TCOC Model excludes an additional 10 specialty care hospitals, children's hospitals, and government-owned hospitals for which HSCRC does not set Medicare payment rates (though they might set rates for other payers).

Exhibit A.1. Hospitals that participated in the MD TCOC Model in 2022—overall and in specific quality programs

Hospital name	Type of hospital (as of 2022)	Hospital rates in RY2022 (July 2021 to June 2022) were adjusted based on performance in the following quality programs:			
		Maryland Hospital-Acquired Conditions Program	Potentially Avoidable Utilization Savings Policy	Quality-Based Reimbursement Program	Readmissions Reduction Incentive Program
Adventist HealthCare Fort Washington Medical Center	Acute-General	Yes	Yes	Yes	Yes
Adventist HealthCare Shady Grove Medical Center	Acute-General	No	Yes	Yes	Yes
Adventist HealthCare White Oak Medical Center (formerly Washington Adventist)	Acute-General	Yes	Yes	Yes	Yes
Adventist HealthCare Germantown Emergency Center	Freestanding Emergency Center or Medical Facility	No	No	No	No
Anne Arundel Medical Center	Acute-General	Yes	Yes	Yes	Yes
Atlantic General Hospital	Acute-General	Yes	Yes	Yes	Yes
CalvertHealth Medical Center	Acute-General	No	Yes	Yes	Yes
Carroll Hospital Center	Acute-General	Yes	Yes	Yes	Yes
Doctors Community Hospital	Acute-General	Yes	Yes	Yes	Yes
Frederick Health Hospital, Inc.	Acute-General	Yes	Yes	Yes	Yes
Garrett Regional Medical Center	Acute-General	Yes	Yes	Yes	Yes
Grace Medical Center (formerly Bon Secours Hospital)	Freestanding Emergency Center or Medical Facility	No	No	No	No
Greater Baltimore Medical Center	Acute-General	Yes	Yes	Yes	Yes
Holy Cross Germantown Hospital	Acute-General	Yes	Yes	Yes	Yes
Holy Cross Hospital	Acute-General	Yes	Yes	Yes	Yes
Howard County General Hospital	Acute-General	No	Yes	Yes	Yes
Johns Hopkins Bayview Medical Center	Acute-General	Yes	Yes	Yes	Yes

Hospital name	Type of hospital (as of 2022)	Hospital rates in RY2022 (July 2021 to June 2022) were adjusted based on performance in the following quality programs:			
		Maryland Hospital-Acquired Conditions Program	Potentially Avoidable Utilization Savings Policy	Quality-Based Reimbursement Program	Readmissions Reduction Incentive Program
Levindale Hebrew Geriatric Center and Hospital	Acute-General	No	Yes	No	Yes
MedStar Franklin Square Medical Center	Acute-General	Yes	Yes	Yes	Yes
MedStar Good Samaritan Hospital	Acute-General	Yes	Yes	Yes	Yes
MedStar Harbor Hospital	Acute-General	Yes	Yes	Yes	Yes
MedStar Montgomery Medical Center	Acute-General	Yes	Yes	Yes	Yes
MedStar Saint Mary's Hospital	Acute-General	Yes	Yes	Yes	Yes
MedStar Southern Maryland Hospital Center	Acute-General	No	Yes	Yes	Yes
MedStar Union Memorial Hospital	Acute-General	Yes	Yes	Yes	Yes
Mercy Medical Center, Inc.	Acute-General	Yes	Yes	Yes	Yes
Meritus Medical Center	Acute-General	No	Yes	Yes	Yes
Northwest Hospital	Acute-General	Yes	Yes	Yes	Yes
Saint Agnes Hospital	Acute-General	Yes	Yes	Yes	Yes
Sinai Hospital	Acute-General	Yes	Yes	Yes	Yes
Suburban Hospital	Acute-General	Yes	Yes	Yes	Yes
The Johns Hopkins Hospital	Acute-General	Yes	Yes	Yes	Yes
The University of Maryland Upper Chesapeake Medical Center	Acute-General	Yes	Yes	Yes	Yes
TidalHealth McCready Pavilion	Freestanding Emergency Center or Medical Facility	No	No	No	No
TidalHealth Peninsula Regional, Inc.	Acute-General	Yes	Yes	Yes	Yes
Union Hospital of Cecil County	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Baltimore Washington Medical Center	Acute-General	Yes	Yes	Yes	Yes

Hospital name	Type of hospital (as of 2022)	Hospital rates in RY2022 (July 2021 to June 2022) were adjusted based on performance in the following quality programs: Maryland Hospital-Acquired Conditions Program	Potentially Avoidable Utilization Savings Policy	Quality-Based Reimbursement Program	Readmissions Reduction Incentive Program
University of Maryland Bowie Health Center	Freestanding Emergency Center or Medical Facility	No	No	No	No
University of Maryland Charles Regional Medical Center	Acute-General	No	Yes	Yes	Yes
University of Maryland Harford Memorial Hospital	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Laurel Regional Hospital	Freestanding Emergency Center or Medical Facility	No	No	No	No
University of Maryland Medical Center	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Medical Center Midtown Campus	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Prince George's Hospital Center	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Queen Anne’s Freestanding Emergency Center	Freestanding Emergency Center or Medical Facility	No	No	No	No
University of Maryland Rehabilitation & Orthopaedic Institute	Acute-General	Yes	No	No	Yes
University of Maryland Shock Trauma	Special emergency hospital	No	No	No	No
University of Maryland Shore Medical Center at Dorchester	Freestanding Emergency Center or Medical Facility	Yes	Yes	Yes	Yes
University of Maryland Shore Medical Center at Easton	Acute-General	Yes	Yes	Yes	Yes
University of Maryland Shore Regional Health at Chestertown	Acute-General	Yes	Yes	No	Yes

Hospital name	Type of hospital (as of 2022)	Hospital rates in RY2022 (July 2021 to June 2022) were adjusted based on performance in the following quality programs:			
		Maryland Hospital-Acquired Conditions Program	Potentially Avoidable Utilization Savings Policy	Quality-Based Reimbursement Program	Readmissions Reduction Incentive Program
University of Maryland St. Joseph Medical Center	Acute-General	Yes	Yes	Yes	Yes
Western Maryland Regional Medical Center	Acute-General	No	Yes	Yes	Yes
Totals	Acute-General: 44 Freestanding Emergency Center or Medical Facility: 7 Special emergency hospital: 1 All: 52	Yes: 37 No: 15 All: 52	Yes: 44 No: 8 All: 52	Yes: 42 No: 10 All: 52	Yes: 45 No: 7 All: 52

Source: HSCRC 2022 financial reports.

ED = emergency department; HSCRC = Health Services Cost Review Commission; RY = rate year.

A.1.1b. Incentive sizes in 2022

Hospital retained revenue

We received retained revenue from HSCRC for the 52 hospitals found in the MD TCOC State Agreement for rate year (RY) 2022, which covers amounts that all payers paid hospitals for services rendered July 2021 to June 2022. Retained revenue is the difference between a hospital's revenues based on its global budget and the lower revenues it would have earned due to volume reductions under an FFS schedule.

Background. To understand HSCRC's method for calculating retained revenue, some background information on how HSCRC sets global budgets will be helpful. At the start of the state fiscal year, HSCRC provides each hospital with its all-payer budget for the year along with the assumptions that went into the budget. HSCRC divides hospital care into 75 cost centers, each with its own unit of measurement. For example, acute medical-surgical is one of the cost centers, and it is measured in patient days. HSCRC sets each hospital's budget by setting a hospital-specific expected volume for each cost center and a hospital-specific rate that all payers will pay for a unit of service for that cost center in the year for that hospital.[1]

- From 2014 to 2021, HSCRC set the expected volumes based on the hospital's actual volume in 2013 (before the Maryland All-Payer Model began) and updated them for population growth and shifts in care across hospitals—but not for any activities that the hospital used to successfully reduce avoidable hospital volumes.
- HSCRC sets rates based on the prior year rates, with a yearly statewide update for all hospitals to account for inflation and other factors, and hospital-specific changes to rates based on performance on quality measures and other factors.

A hospital's all-payer budget for the year is the sum across all cost centers of the expected volume times the approved rate in that cost center. During the rate year, if the actual hospital volumes are below the volumes built into the budget, the hospital can increase the prices it charges to payers so that by the end of the year, the actual volumes times the actual prices equal the global budget amount.

Starting in RY2022, HSCRC "rebased" volumes—meaning that they set the expected volumes in the budget to be a hospital's actual volume in 2019. To ensure this rebasing did not affect a hospital's global budget amount, HSCRC changed the rates (increasing them for most hospitals) so that the new expected volume (lower for most hospitals) times the new rates would equal their original budget. HSCRC did the rebasing in part so that hospitals would be less likely to exceed the price changes that hospitals can make during the year without asking for special permission from HSCRC. Based on this policy, hospitals typically need to request permission to raise their actual prices more than 5% above the approved rates. However, given historical volume declines,

[1] As a result, the rate for a service is the same for a given hospital across all payers, though the rates differ across hospitals.

many hospitals were at or near this cap. Rebasing meant these hospitals were no longer at the cap because they could charge actual prices much closer to their approved rates.

HSCRC adjusted the retained revenue calculations for this rebasing, by essentially adjusting the 2022 retained revenue to reflect volume reductions that occurred before 2019. Specifically, HSCRC followed the same method described below but adjusted the approved rate (used to calculate the "standard revenue") down if the hospital retained revenue in calendar year 2019—and thus reduced volume before 2019—or up if the hospital had a negative retained revenue value in 2019. The adjustment factor was based on the charge variation in 2019, which is the percentage difference between the actual revenue and the "standard revenue". The charge variation is a positive percentage if the actual revenue is higher (and thus, the hospital retained revenue) and is a negative percentage if the actual revenue is lower. HSCRC divided the approved rate by 1 plus the 2019 charge variation and then used that rate to calculate the "standard revenue" in 2022. If the hospital had retained revenue in 2019, the 2022 retained revenue would increase after this adjustment, reflecting the full amount of revenue the hospital retained in 2022 based on volume reductions since 2013. Conversely, if the hospital did not retain revenue in 2019, the 2022 retained revenue would decrease after this adjustment. Dividing out the prior charge variation rather than adding retained revenue from prior years accounts for inflation by avoiding the use of a dollar scale, which has changed in value over time.

HSCRC's method. The HSCRC method for calculating retained revenue compares a hospital's actual revenue for the year with their "standard revenue." For each hospital, the "standard revenue" is the product of the actual volume for the rate year and the approved rate, summed across the hospital's cost centers. That is, the standard revenue captures the amount that the hospital would have received for the actual volume of care they provided at the price that HSCRC set at the start of the year—not the price that the hospitals actually charged. For a hospital that reduced volume relative to the expectation built into the budget, the approved rate will be below the actual price charged during the year. In this case, actual revenue will be higher than the standard revenue, so retained revenue (the difference between the two) will be positive. Conversely, if a hospital has volumes that are higher than those built into the budget, then their actual revenue will be lower than their standard revenue, and they will have negative retained revenue.

Retained revenue reflects the size of the incentive for hospitals to reduce avoidable volume. Because the expected volumes built into the budgets are based on 2013 volumes, any efforts to successfully reduce volume below that level will generate retained revenue—and the more a hospital reduces volume relative to its 2013 baseline, the larger its retained revenue.

Retained revenue for RY2022. HSCRC provided us with retained revenue for each hospital for RY2022, calculated using the method described above. This includes accounting for the "rebased" volumes as described earlier.

To calculate total retained revenue for RY2022 across the state, we summed all negative retained revenue separately from all positive retained revenue across hospitals for final amounts of

-$79,429,840 across 11 hospitals (negative retained revenue) and $1,116,286,224 across the remaining 41 hospitals (positive retained revenue).

A.1.2. Component 2: Maryland Primary Care Program

A.1.2a. Participation and reach

We obtained Maryland Primary Care Program (MDPCP) participation data from The Lewin Group, the contractor that helps implement MDPCP. The Lewin Group provided a roster of the practices participating in MDPCP, including the practice's Tax Identification Number (TIN), a list of practitioners working at each practice, a list of Medicare fee-for-service (FFS) beneficiaries attributed to each practice, and a list of Care Transformation Organizations (CTOs) that supported MDPCP practices.

To identify the number of practices (and CTOs) that participated in MDPCP in 2022, we limited the sample of practices to those that participated for at least part of that year (508 practices) and the CTOs that supported at least one practice in 2022 (24 CTOs).

We calculated the reach of MDPCP in two ways: the percentage of primary care practices in the state participating in MDPCP in 2022 (practice-level reach), and the percentage of Medicare FFS beneficiaries in the state attributed to MDPCP practices in 2022 (beneficiary-level reach).

Practice-level reach

We used data from The Lewin Group on MDPCP practices and providers and IQVIA's OneKey™ data to calculate the reach of MDPCP among primary care practices over time. The OneKey™ data contains information about practices throughout the United States, including the practices' location, specialty, the providers affiliated with the practice, and the corporate parents of the practices. We conducted the following steps to merge the Lewin Group and OneKey™ data and identify which practices in Maryland were eligible for MDPCP (defined as having at least one primary care provider and at least 125 attributed Medicare FFS beneficiaries):

1. First, we limited the OneKey™ data to practices in Maryland and identified the MDPCP practices that ever participated in MDPCP from 2019 to 2022 in the 2019 to 2022 OneKey™ data. To link the two data sets, we merged the National Provider Identifiers (NPI) of the affiliated providers at the practices and retained all practices that were found in both data sets. We compared practice zip code, address, and name (using Levenshtein difference in characters) among these overlapping practices. Matches were confirmed if all three variables matched across the two data sets; otherwise, the matches were adjudicated by two independent reviewers.
2. Second, we determined whether practices had an eligible primary care provider by linking specialty information from the National Plan and Provider Enumeration System to the list of affiliated providers in the OneKey™ data.
3. Third, we assigned a Tax Identification Number (TIN) to all Maryland practices in the OneKey™ data and implemented attribution at the TIN-NPI level based on the MDPCP attribution approach (see Appendix E for full details on TIN assignment and attribution).

To calculate the numerator for the practice-level reach, we used the count of unique MDPCP practices participating in each year (2019–2022) [N = 508 in 2022]. To calculate the denominator, we added the number of all MDPCP practices from the numerator with the number of all otherwise eligible non-MDPCP practices in the OneKey™ data in Maryland in each year. In 2022, there were 1,472 unique practices in the OneKey™ data in Maryland. We consider these practices ineligible for MDPCP if they: (1) did not have at least 125 attributed beneficiaries, including those that did not have attributed beneficiaries because they did not have at least one eligible primary care provider or an assigned TIN (N = 1,174), and (2) were not either primary care practices (including family or general practice, geriatrics, internal or preventive medicine, pediatric medicine, primary care or multispecialty group practices based on specialty information in OneKey™ or were affiliated with the Veteran's Health Administration or a health department (N = 27). This resulted in 271 non-MDPCP practices in Maryland that were otherwise eligible for MDPCP in 2022. The 2022 denominator was the 779 (508 MDPCP practices plus the 271 non-MDPCP practices that were deemed eligible). The practice-level reach was 65% in 2022 (N = 508/779).

Beneficiary-level reach

We used data from The Lewin Group on attributed MDPCP beneficiaries and information from Medicare FFS enrollment data to determine the beneficiary-level reach of MDPCP over time among Medicare FFS beneficiaries in Maryland. We took the list of MDPCP beneficiaries in 2022 (N = 429,943) from The Lewin Group, de-duplicated it (N = 409,209), and limited it to beneficiaries that reside in Maryland and are enrolled in Part A and B coverage for at least one month in 2022 (N = 400,818) to identify the numerator. We used the count of all Medicare FFS beneficiaries with Part A and B coverage enrolled in Maryland who are enrolled for at least one month in 2022 by Chronic Conditions Warehouse Beneficiary ID (N = 777,372). The beneficiary reach was 52% in 2022 (400,818/777,372).

A.1.2b. Payments to practices and CTOs in 2022

We obtained MDPCP payment data from The Lewin Group. This data set includes CMS payments to practices and CTOs from four payment streams: care management fees, performance-based incentive payments, Health Equity Advance Resource and Transformation (HEART) payments, and comprehensive primary care payments. When describing total CMS payments in calendar year 2022, we summed 2022 payments provided to the 508 practices that ever participated in MDPCP in 2022. The following is a description of each of the payment streams (CMS 2021a). (Exhibit A.2 shows the totals that CMS paid for each type in 2022, for practices and CTOs.)

- Care management fees (CMFs) support practices to make care delivery changes. Track 2 practices received higher CMFs to meet the more intensive care transformation requirements. For both tracks, CMFs are adjusted based on beneficiary risk tiers and are not tied to practice performance. If practices elect to partner with a CTO, they share 30% or 50% of their CMF with their CTO partner.

- Performance-based incentive payments (PBIPs) are prospectively paid at $2.50 per beneficiary per month for Track 1 practices and $4.00 for Track 2 practices. Based on clinical quality and utilization performance measures, CMS could partially or entirely recoup the payments. If practices received a score of 50% or less, CMS would recoup the entire PBIP, and if practices scored 80% or higher, the practice would keep the entire amount. CTOs also receive PBIPs at the start of each performance year at a rate of $4.00 per beneficiary per month, and are also subject to recoupment based on the performance of practices they serve. In 2021, CMS recouped about 40% of the PBIP payments (recoupment data for 2022 not available at the time we wrote this appendix).
- HEART payments are new starting in 2022, and they are paid quarterly for beneficiaries with both medical complexity and high measured social disadvantage (CMS 2022a). Medical complexity is defined as at or greater than 75th percentile of Hierarchical Condition Category risk scores based on all Maryland beneficiaries, and measured social disadvantage is defined as the top quintile of the Area Deprivation Index among the MDPCP beneficiary population (Center for Health Disparities Research n.d.).[2] Practices are paid $110 per beneficiary per month for eligible beneficiaries.
- Comprehensive primary care payments are provided to Track 2 practices to help enable flexible coordination of care. They are a hybrid alternative to FFS payments, where CMS prospectively pays practices a percentage of expected Medicare payments for evaluation and management services. Track 2 practices can elect to receive 10, 25, 40, or 65 percent of payments through the comprehensive primary care payments instead of FFS payments and are expected to increase their percentage over time.

[2] Hierarchical Condition Category risk score tiers are based on percentile distributions based on all beneficiaries in Maryland specified in the MDPCP Participation Agreements. Area Deprivation Index quintiles are based on a distribution of all attributed MDPCP beneficiaries, updated quarterly.

Exhibit A.2. Total CMS payments in 2022 to MDPCP practices and CTOs, by type of payment and overall

Type of payment	CMS payments to all MDPCP practices	CMS payments to all CTOs
Number	508 practices total (500 received HEART payments)	24 CTOs
Enhanced payments (% of total enhanced payments)		
Care management fee	$78 million (70%)	$36 million (63%)
HEART payments	$19 million (17%)	$12 million (21%)
PBIP	$14 million (13%)	$9 million (16%)
Total enhanced payments	$111 million	$57 million
Comprehensive Primary Care Payments (which replace some portion of FFS payments)	$27 million	--
Total	**$138 million**	**$57 million**

Source: Lewin 2022 monitoring data, current as of 2023 quarter 1.

Note: PBIP is 2022 payments before recoupment. In 2021, CMS recouped 40% of practices' and CTOs' PBIP based on quality, utilization, and efficiency.

CMS = Centers for Medicare & Medicaid Services; CTO = Care Transformation Organization; FFS = fee for service; MDPCP = Maryland Primary Care Program; PBIP = performance-based incentive payment.

To calculate MDPCP payments as a share of practice revenue, we obtained self-reported data on practice revenue from the CMS MDPCP portal. Out of the 496 practices that participated through the end of 2022,[3] 494 self-reported their revenue for calendar year 2022. For each of those 494 practices, we calculated enhanced payments (including the CMF, PBIP, and HEART payments) as a percentage of the practice's total self-reported revenue. We then took the median of this percentage across the 494 practices to conclude that MDPCP increased practice revenue by about 10%.

A.1.3. Component 3: Quality adjustments to hospital global budgets

A.1.3a. Participation and reach

For most of the 52 hospitals participating in the MD TCOC Model, HSCRC adjusted the hospital's budget for rate year (RY) 2022[4] based on the hospital's measured performance on quality measures. HSCRC adjusted payments in RY2022 for four quality programs (Exhibit A.3). Some hospitals did not participate in specific quality programs in 2022 because they were not eligible. For example, free-standing emergency departments are not eligible for the readmissions incentive program because they do not have any inpatient admissions that would trigger as possible readmission. Similarly, some hospitals are not eligible for the Hospital-Acquired Conditions Program because they do not have enough at-risk admissions to reliably assess a hospital's performance.

[3] Although 508 practices were participating at the start of 2022, three practices merged with other practices and nine practices dropped out during the year, leaving 496 that participated through the end of 2022.

[4] Rate Year 2022 ran from July 2021 to June 2022.

Exhibit A.1 indicates which quality programs each hospital participated in for RY2022 and the total number of hospitals that participated in each program (bottom row). We developed this list by downloading spreadsheets from HSCRC that describe quality payments, by hospital, in RY2022 for each of the four quality programs (Exhibit A.3 includes links to the data sets). We merged those spreadsheets with the list of 52 hospitals participating in the MD TCOC Model. If a hospital did not appear in the spreadsheet for the quality payments in RY2022, we marked the hospital as not participating in that quality program.

Exhibit A.3. HSCRC's four hospital quality programs and location of data sets for participation and incentives in each one

Quality program	Description	Example measures	Link for data set with hospital participation and incentives in RY2022 (accessed 08/14/2023)
Maryland Hospital-Acquired Conditions Program	Adjusts hospital budgets based on their performance on 14 identified potentially preventable complications.	Recalibrated Patient Safety Indicator measure; Central Line-Associated Bloodstream Infection	https://hscrc.maryland.gov/Documents/Hospitals/gbr-tpr-update/FY-2022/RY2022MHAC.xlsx
Potentially Avoidable Utilization Savings Policy	Applies penalties for performance on per-capita potentially avoidable admission rates.	PQI 90 Prevention Quality Overall Composite; PQI 93 Prevention Quality Diabetes Composite	https://hscrc.maryland.gov/Documents/Hospitals/gbr-tpr-update/FY-2022/FINAL_RY2022_PAU_Savings_Scaling.xlsx
Quality-Based Reimbursement Program	Incentivizes quality improvement across domains of person and community engagement, clinical care, and patient safety.	Communication with Nurses; Cleanliness and Quietness of Hospital Environment; All Condition Inpatient Mortality	https://hscrc.maryland.gov/Documents/Hospitals/gbr-tpr-update/FY-2022/RY2022QBR.xlsx
Readmissions Reduction Incentive Program	Incentivizes hospitals to reduce avoidable admissions within 30 days of discharge from the hospital.	30-day all-payer, all hospital readmission rate with case-mix adjustments	https://hscrc.maryland.gov/Documents/Hospitals/gbr-tpr-update/FY-2022/RY2022RRIP.xlsx

HSCRC = Health Services Cost Review Commission; PQI = Prevention Quality Indicators; RY = rate year.

A.1.3b. Incentive sizes in 2022

We obtained data on each hospital's quality adjustments to global budgets for RY2022 from the HSCRC website (HSCRC n.d.) (see Exhibit A.3). These adjustments reflect the total amount that each hospital earned or lost based on performance on quality measures in HSCRC's various quality programs. HSCRC made these adjustments to global budgets by increasing the all-payer rates (and therefore the global budget) for the higher-performing hospitals and decreasing them for the lower-performing hospitals.[5] These adjustments were based on calendar year 2016 benchmarks and measured performance in calendar year 2019. HSCRC would typically have used a more recent performance year (2020) to reward or penalize hospitals in RY 2022, but they used 2019 so that care pattern changes during the COVID-19 pandemic would not interfere with the quality measures.

After downloading the incentive payments for each quality program and each hospital in RY2022, we summed the negative and positive amounts separately across hospitals for each quality program (see Exhibit A.4). Finally, we summed the total negative and positive amounts across the four quality programs to calculate the total negative quality adjustments (-$102,424,460) and total positive quality adjustments ($85,623,635) provided to hospitals.

Exhibit A.4. Total positive and negative adjustments in RY2022 across participating hospitals

Quality program	Number of participating hospitals	Total negative amount	Total positive amount
Maryland Hospital-Acquired Conditions Program	37	-$3,345,082	$43,255,352
Potentially Avoidable Utilization Savings Policy	44	-$39,559,507	$0
Quality-Based Reimbursement Program	42	-$53,864,068	$2,820,647
Readmissions Reduction Incentive Program	45	-$5,655,803	$39,547,636
All	45[a]	-$102,424,460	$85,623,635

Source: HSCRC financial reports.

HSCRC = Health Services Cost Review Commission; RY = rate year.

[a]45 hospitals participated in at least one of the four quality programs in RY2022.

A.2. Components with mid-range incentives or investments in 2022 ($10 million to $100 million)

A.2.1. Component 4: Medicare Performance Adjustment

The Medicare Performance Adjustment (MPA), which adjusts each hospital's budget based on the total Medicare spending for beneficiaries attributed to them, places some accountability on hospitals for total cost of care (TCOC).

We obtained the Year 4 Medicare Performance Adjustment Monitoring Report from the CRISP website, which contains data on the traditional MPA and adjustments by hospital for RY2023

[5] All of the quality adjustments are two sided (meaning that hospitals could gain or lose under the adjustment) except for the Potentially Avoidable Utilization Savings policy adjustment. That adjustment is one-sided; hospitals can only lose or have no net change in their budget based on their Potentially Avoidable Utilization performance.

(which runs from July 2022 to June 2023). The data are based on calendar year 2021 performance compared with a benchmark of calendar year 2019. The data set is located here: https://www.crisphealth.org/wp-content/uploads/2022/07/Final-Y4-MPA-CMMI-6-1.xlsx, accessed August 15, 2023.

For RY2023, there were two components to the MPA. The first component, the traditional MPA, adjusted a hospital's budget based on the TCOC performance for all attributed beneficiaries, with attribution based on whether beneficiaries lived within the hospital's service area. This adjustment increased or decreased a hospital's budget by up to 1% of their Medicare revenue. The second component, the MDPCP supplement, further adjusted hospital budgets for the TCOC experience of beneficiaries attributed to them through the hospital role as CTO supporting primary care transformation. HSCRC included this supplement to provide hospitals, or their systems, with additional accountability for TCOC performance for beneficiaries served by MDPCP practices that the hospital supported in its role (or its system's role) as a CTO.

In RY2023, 44 of the 52 hospitals received an adjustment based on the traditional MPA. The hospitals that did not have an adjustment were not eligible because they were a free-standing emergency room. In all, 37 of the 52 hospitals also received the MDPCP supplemental adjustment because they, or their system, served as a CTO for MDPCP practices.

To calculate the statewide MPA incentive in RY2023, we summed the negative and positive traditional MPA and MDPCP supplemental adjustment amounts separately across hospitals (see Exhibit A.5). Then, we combined the total negative and positive amounts across programs to calculate the MPA amount including the MDPCP supplemental adjustment, which is a sum of $24,995,890 in positive adjustments and -$41,504,886 in negative adjustments.

Exhibit A.5. Adjustments to hospital budgets through the MPA

Type of MPA adjustment	Goal	Sum of the positive adjustments (# of hospitals in the sum)	Sum of the negative adjustments (# of hospitals in the sum)
Traditional	To place some TCOC on all hospitals in the MD TCOC Model	$12,273,597 (N = 24)	-$21,530,782 (N = 20)
MDPCP supplement	To place additional accountability on hospitals (or their systems) that function as CTOs in the MDPCP program	$12,722,293 (N = 19)	-$19,974,104 (N = 18)
Total	--	**$24,995,890**	**-$41,504,886**

Source: HSCRC Year 4 MPA Monitoring Report.

CTO = Care Transformation Organization; MD TCOC = Maryland Total Cost of Care; MDPCP = Maryland Primary Care Program; MPA = Medicare Performance Adjustment; TCOC = total cost of care.

A.2.2. Component 5: Regional Partnership Catalyst Grants

HSCRC created the Regional Partnership Catalyst Program to help hospitals achieve two of the state's population health goals under the MD TCOC Model: (1) reduction in statewide BMI (relative to comparison group), which is linked to diabetes risk, (2) reduction in overdose deaths (also relative to a separate comparison group; HSCRC 2023). HSCRC provided five-year grants to hospitals, which would work together and with community partners to implement interventions in service of either one or both of these goals.

We downloaded data from the HSCRC for the amounts that each hospital received in Regional Partnership Catalyst Grant funding for RY2022: https://hscrc.maryland.gov/Pages/hsp-gbr-tpr-update.aspx, accessed July 1, 2022. Based on this data set, 33 hospitals received Regional Partnership funding for one or both the catalyst grants in RY2022. We summed the funding amounts across hospitals as an additional incentive under the MD TCOC Model, at a total of $26,854,323. Although the grants total are over five years, we used the amount that hospitals received in just RY2022 to put them on the same yearly scale as the other incentives and supports we report.

A.3. Components with relatively small or no incentive payments in 2022 (less than $10 million)

A.3.1. Components 6 and 7: Hospital-operated episode programs

HSCRC also adjusts the amounts that hospitals receive each year based on their performance in two episode programs specific to Medicare FFS beneficiaries: Episode Care Improvement Program (ECIP) and Care Transformation Initiatives (CTIs). In each program, HSCRC measures actual Medicare Part A and B spending for episodes against a benchmark. If spending falls below the benchmark, hospitals can earn an incentive payment. Both programs are voluntary in the sense that hospitals choose whether to participate. For CTIs, however, HSCRC reduces Medicare prices across all hospitals to fund the payments to hospitals that earn incentives. In this way, CTI payments are cost neutral to CMS (they shift Medicare spending across hospitals in Maryland but do not increase Medicare spending).

HSCRC makes payments to hospitals that are successful in ECIP and CTIs through the MPA. Although HSCRC sets prospective rates that are the same across all payers at the start of each fiscal year, HSCRC uses the MPA to change the actual prices that Medicare pays relative to other payers. For example, for a hospital that is successful under ECIP or CTIs, HSCRC can direct Medicare to pay that hospital 1% higher Medicare prices than the prospectively set rates. Each year, HSCRC produces an annual MPA Monitoring Report that includes how much each hospital earned in ECIP and CTIs and, accordingly, how much they adjusted the prices that Medicare paid hospitals to cover the earned amounts.

- **Component 6: Episode Care Improvement Program**
  - *Participation.* We developed a list of 2022 ECIP participants by compiling implementation plans from Salesforce and cross-referencing the participants with the list on HSCRC's website: https://hscrc.maryland.gov/pages/careredesign.aspx, accessed June 21, 2022. 24 hospitals participated in ECIP in 2022.
  - *Reach.* We calculated the reach of ECIP at the hospital discharge level because ECIP is generally focused on distinct episodes of care triggered by a hospital discharge. We received data on the number of ECIP episodes in 2022 from HSCRC. To calculate the total number of discharges in Maryland in 2022, we summed the total number of all-cause admissions among Maryland FFS beneficiaries (based on Medicare Part A claims). Dividing the total number of ECIP episodes by the total number of discharges in Maryland gives us an ECIP reach of 3.6% of discharges in 2022 (=5,978/164,075).
  - *Payments.* We used the HSCRC Year 4 MPA Monitoring Report to calculate total ECIP payments for RY2023 (which ran from July 2022 to June 2023). ECIP payments are provided biannually: once in January and once in July. Preliminary payments are provided one year after the performance period concludes, and then adjustments are made six months later. For this incentive, we included the preliminary payments made to hospitals in January 2021 and adjustments made in July 2021 (reflecting performance from July to December 2020) as well as the preliminary payments reflecting performance from January 2021 to June 2021. These payments are applied to hospital budgets for RY2023.
- **Component 7: Care Transformation Initiatives**
  - *Participation.* We obtained a list of 2022 CTI participants and their CTI selections from the Care Transformation Profiler in the CRISP Reporting Services Portal. In total, 42 hospitals participated in one or more CTIs in 2022.
  - *Reach.* Although many CTIs begin with a discharge, a substantial number of CTIs are in the primary care thematic area, which encompasses attributed beneficiaries that may or may not use hospital services during the year. For this reason, we calculated the reach of CTIs at the beneficiary level. We received data from HSCRC on the deduplicated number of beneficiaries participating in CTIs in 2022 and calculated the total number of beneficiaries in Maryland using Medicare claims data.[6] Dividing the number of distinct beneficiaries under CTIs by the total number of Medicare FFS beneficiaries in Maryland gives us a CTI reach of 28.4% of beneficiaries in 2022 (=220,823/777,372).
  - *Payments.* The first measured performance period for CTIs was from July 2021 to June 2022. The first payments were not made until RY2024—in July 2023—which is outside the time period covered for the impact analysis in this report. However, with a reach of almost one-third of beneficiaries, payments in RY2024 were substantial. Specifically,

[6] For the denominator, we used the 777,372 Medicare FFS beneficiaries who are in the Maryland analytic sample for impact analysis in 2022. This limits it to Medicare FFS beneficiaries with both Part A and B coverage, who had Medicare as their primary payer of medical bills, and who were not enrolled in a Medicare Advantage plan.

after adjusting for quality, 16 hospitals earned positive payments totaling $56,316,813. In contrast, 28 hospitals had negative adjustments totaling $56,142,388. These negative adjustments generated the revenue needed to pay the positive adjustment to the successful hospitals while remaining cost neutral to Medicare.

A.3.2. Component 8: Episode Quality Improvement Program

- *Reach and payments.* HSCRC reported in its January 2022 Episode Quality Improvement Program (EQIP) subgroup meeting that there were 1,979 Care Partners (specialists) participating in EQIP in 2022 in the orthopedics, cardiology, and gastroenterology clinical episode categories (HSCRC 2022a). Similar to CTIs, the first payments were not made until RY2024. As the CRP Entity, University of Maryland Medical Center received $10.8 million in incentive payments and then distributed the payments to each EQIP Entity (individuals or groups of specialists) based on performance. Each EQIP Entity can determine how to distribute the payments among each Care Partner.

A.4. Comparison of measures in this report to those in the state agreement and SIHIS

The 13 outcomes included in the Progress Report for Medicare impact analyses (Chapter 2, 4, and 5) and the 5 outcomes included in Medicaid trends analyses (Chapter 3) partially align with those in the legal agreement that established the MD TCOC Model (Exhibit A.6) and SIHIS (Exhibit A.7). The Progress Report does not contain all the state agreement or SIHS-related measures because of data limitations or because they were out of the scope of the evaluation. Further, although some of the state agreement and SIHIS measures focus on all Maryland residents, the measures in Chapters 2, 4 and 5 are for Medicare fee-for-service (FFS) beneficiaries, with the exception of patients' ratings: for hospitals, which is measured among all patients, and for their personal doctor, which is measured among all Medicare beneficiaries (FFS and Medicare Advantage). Chapter 3 describes trends over time for a smaller set of measures for Maryland Medicaid and Children's Health Insurance Program (CHIP) enrollees.

Although we aligned outcome measures in the Progress Report with the state agreement and SIHIS when feasible and appropriate, we did not aim to align methods for estimating effects on Medicare FFS beneficiaries for these measures. The state agreement and SIHIS set their own methods for assessing progress toward the stated goals, which typically do not rely on a matched comparison group. By contrast, all the Medicare impact estimates in the report use difference-in-differences models with matched comparison groups.

The Medicare FFS impact analyses include eight outcome measures for Medicare FFS beneficiaries that are not explicit goals in either the state agreement or in SIHIS. We included these measures in the report because the model's incentives and supports could logically lead to improvements in them or because the model could have unintended consequences, worsening these outcomes. The outcome measures are the following:

1. All-cause acute care hospital admissions
2. Outpatient ED visits and observations of stays
3. Intensity of hospital care (measured by standardized hospital spending)
4. Non-hospital spending
5. Post-acute care spending
6. Patients' rating of their personal doctor
7. Patients' rating of their hospital

The Medicaid/CHIP trend analyses include the first two measures above – all-cause hospital admission and outpatient ED visits and observation stays. We analyzed these measures for the Medicaid/CHIP population for same reasons described above for the Medicare analyses – that is, the model's global budgets, which are all-payer, and related incentives to reduce unnecessary hospital use, likely influenced Medicaid/CHIP trends. We also included a measure of total spending in the Medicaid/CHIP trend analyses, defined as total FFS payments plus total capitated payments made by the state to managed care organizations. Ideally, we would have assessed trends in hospital versus non-hospital spending for Medicaid/CHIP. However, we were unable to break out spending for hospital versus non-hospital services; nearly all enrollees were covered by managed care plans, and we only had data on capitated payments made to managed care plans (and not amounts paid by managed care plans for covered services).

Exhibit A.6. Alignment between outcome measures in the state agreement and the Medicare impact analyses in the Progress Report (Chapters 2, 4, and 5)

Category	Measure	Commitment in the state agreement	Similar measure in chapters 2, 4, and 5 estimates of model impacts
Spending	All-payer hospital spending per Maryland resident	Limit growth to no more than 3.58% per year (the long-term growth rate of the state economy)	Medicare FFS spending for hospital care per Maryland Medicare beneficiary per year
	Annual Medicare FFS savings	Meet specific annual savings targets—for example, $300 million in 2023 (assessed by comparing actual Medicare spending in Maryland with what spending would be if Maryland's 2013 spending grew at the national rate). These annual savings sum to about $2 billion over 8 years (2019 to 2026). Meet guardrail tests: The growth rate in per capita Medicare FFS spending in Maryland must not exceed that in the nation by more than 1 percent in any model year and must not exceed that in the nation by any amount for two or more consecutive years.	Total Medicare FFS spending per year (with and without non-claims payment)
Quality of care	Medicare readmissions rate	Must at least maintain improvements achieved during MDAPM	Medicare 30-day post-discharge unplanned readmission rate
	All-payer reductions in hospital-acquired conditions	Must at least maintain improvements achieved during MDAPM	None

Source: HSCRC 2018.

FFS = fee for service; MDAPM = Maryland All-Payer Model

Exhibit A.7. Alignment between outcome measures and populations in SIHIS with the measures and populations in the Medicare impact analyses (Chapters 2, 4, and 5) and Medicaid/CHIP trends analyses (Chapter 3) in the Progress Report

SIHIS outcome measures, baseline performance, and targets			Impact estimates for Medicare FFS (Chapters 2, 4, and 5) or trends analyses for Medicaid/CHIP (Chapter 3)	
Outcome measure (population)	**2018 baseline**	**2026 final target**	**Related outcome measure**	**Analysis population(s)**
SIHIS domain: Hospital quality				
Avoidable Admissions: Risk-Adjusted PQI-90 Rates (all payer)	1,324 admits per 100,000	25% improvement	Potentially preventable admissions (using the PQI-90 composite)	Medicare FFS Medicaid/CHIP
Readmission disparities (all payer)	Hospital-specific risk difference across levels of Patient Adversity Index	Half of eligible hospitals achieving 50% improvement in disparity	30-day unplanned readmission rates (by race and Social Vulnerability Index)	Medicare FFS
SIHIS domain: Care transformation across the system[a]				
Timely Follow-up After Acute Exacerbations of Chronic Conditions (Medicare FFS)	70.85%	75.00% (5.9% improvement)	Timely follow-up after acute exacerbations of chronic conditions	Medicare FFS
SIHIS domain: Total population health				
Mean BMI in the population of adult Maryland residents (all residents)	28.13 kg/m2	Achieve more favorable change from baseline compared with control	Use of Diabetes Prevention Program services [an intermediate goal in SIHIS]	Medicare FFS
Overdose mortality (all residents)	Age-adjusted death rate of 37.2/100,000	Achieve more favorable change from baseline compared to control	None	n.a.
SMM (all payer)	243.1 SMM rate per 10,000 delivery hospitalizations	197.1 SMM rate per 10,000 delivery hospitalizations	None	n.a.
Asthma-related ED visit rate for ages 2-17 (all payer)	9.2 visits per 1,000	5.3 per 1,000	Asthma-related ED visits for ages 2-17	Medicaid/CHIP

Source: HSCRC 2023.

[a] SIHIS has one other measure in this domain: “Increase the amount of Medicare Total Cost of Care or number of Medicare beneficiaries under Care Transformation Initiatives, Care Redesign Program, or successor payment model.” This is a process measure to show how the MD TCOC Model is being implemented over time—and the reach of alternative payment approaches within Maryland. But because it’s not a quality or efficiency outcome for individual people, we are not estimating impacts on this measure.

BMI = body mass index; CHIP = Children’s Health Insurance Program; ED = emergency department; FFS = fee for service; n.a. = not applicable; PQI = Prevention Quality Indicators; SIHIS = Statewide Integrated Health Improvement Strategy; SMM = Severe Maternal Morbidity.

Appendix B. Methods and Supplemental Results for Estimating Statewide Model Effects for Medicare Fee-For-Service Beneficiaries

This appendix provides details and results related to the estimation of the Maryland Model's statewide effects and the possible drivers of those effects. The first section (B.1.) covers detailed methods for estimating statewide effects as well as supplemental results. The second section (B.2.) describes the measures we used and how we define them—including several measures used in other chapters for estimating the effects of the model on health equity or the added effect of the Maryland Primary Care Program (MDPCP). Finally, the last section (B.3.) describes other methods and findings for key drivers of statewide effects (specifically, the responses from hospitals we gathered from hospital site visits and our hospital survey). See Appendix E for details on MDPCP, another possible driver of statewide effects.

B.1. Detailed methods for estimating impacts and supplemental results

B.1.1. Design for estimating impacts

We used a difference-in-differences analysis to estimate the impacts of the model on utilization, spending, quality of care, and population heath—for Medicare fee-for-service (FFS) beneficiaries throughout Maryland from 2014 to 2022. The difference-in-differences framework estimates impacts by comparing changes in outcomes over time for Medicare beneficiaries in Maryland with contemporaneous changes for a similar comparison group selected from outside Maryland. To select the comparison group, we used areas from across the nation, weighted to look like Maryland on many dimensions (including baseline levels and trends in key outcomes) so that the core assumption behind the difference-in-differences model is credible. That assumption is that the changes in outcomes for the comparison group accurately reflect the changes that would have occurred in Maryland absent the model. We needed to draw the comparison group from outside Maryland because the model is statewide with the potential to affect everyone in the state. Regression models improve the precision of the estimates and adjust for any observed differences between Maryland and the matched comparison group. The regression analyses use different units of analysis depending on the outcome (for example, the unit is the beneficiary-year for outcomes measured at the beneficiary level and discharge-year for those measured at the discharge level).

We are using 2011 to 2013 as the baseline period because doing so permits us to estimate impacts of the model by year. By matching Maryland to a comparison group with similar outcome trends from 2011 to 2013, we aimed for the comparison group to reflect the path that Maryland would have been on if it had not introduced any of the changes starting in 2014—the counterfactual. These changes include the hospital global budgets that started with the Maryland All-Payer Model (MDAPM) in 2014, and the broader state accountability for cost and quality of care and corresponding broadening of incentives to providers that began with MD TCOC in 2019. Using this comparison group, we can directly estimate the accumulated effects of all changes since 2014. We can then use these yearly estimates to combine and compare estimates across time periods. For example, we can compare the average effects during the MD TCOC period (2019–2022) to the effects at the end of the MDAPM period (2017 and 2018) to comment on whether, and how much, effects have grown since the start of the MD TCOC period. For this report, we focused on this specific contrast—the MD TCOC period versus the end of the MDAPM period—because we think it represents a meaningful comparison, but the flexible yearly effect estimates approach allows the reader to make other comparisons as needed.

Because MDAPM and MD TCOC are statewide initiatives, the evaluation ultimately aims to measure population-level impacts for Maryland's entire Medicare FFS population. Thus, our primary impact analyses apply the difference-in-differences design to repeated cross sections of all observable Medicare FFS beneficiaries living in Maryland in each year.[7] The analytic file covers a pre-intervention period three years before MDAPM began (2011 to 2013), the MDAPM period (2014 to 2018), and a period after MD TCOC was implemented in 2019 and ending in 2022 for this report.

B.1.2. Developing the matched comparison group

We developed the matched comparison group in four steps:

1. Selected the unit of analysis for matching
2. Identified variables to match on and set criteria for what counts as sufficient balance
3. Used a reweighting method to create the matched comparison group
4. Assessed the quality of the matched comparison group in terms of balance, size, geographic spread, and statistical power

In the following sections, we describe each of these four steps. When we developed the comparison group, we explored many alternatives reflecting tradeoffs in different dimensions of quality for the comparison group. We discussed these alternatives with CMS and decided on a final comparison group that we agreed achieved the best balance on the various dimensions. In this section, we report only the results for the final selected comparison group.

B.1.2a. Selecting the unit of analysis for matching

We selected Public Use Microdata Areas (PUMAs) as the unit for matching. PUMAs are large enough to limit variation in outcomes attributable to random noise but small enough to capture meaningful variation within populous and diverse counties. Specifically, there are 44 PUMAs in Maryland, and the potential comparison group included 2,336 PUMAs from the remaining 49 states plus Washington, DC. PUMAs, defined by the U.S. Census Bureau, are built on census

[7] We define a beneficiary as observable in the year if they are alive, enrolled in FFS Medicare with Part A and B, and have Medicare as the primary payer in at least one month of the year. We allow beneficiaries to be observable for only part of the year (as little as a single month based on meeting the criteria above). In those partial year observability cases, we annualize outcomes (projecting what outcomes would have been over a full year) and then weight by observability in the regressions, down-weighting beneficiaries who are observed for less than a full year proportional to the amount of time we observe them.

tracts and counties and contain at least 100,000 people. Larger counties such as Baltimore City (a county equivalent) are divided into multiple PUMAs, enabling finer resolutions for determining whether key contextual factors vary within the county. Sparsely populated counties are combined into a single PUMA to help ensure that any statistics calculated for this population are reliable.

B.1.2b. Identifying variables to match on and setting criteria for what counts as sufficient balance

In close collaboration with CMS, we set priorities for matching variables to make the matching process feasible and on target (summarized in Exhibit B.1). So that the matched comparison group would estimate Maryland's counterfactual, we set out to select a comparison group that had the following:

- Parallel trends for priority outcomes during the baseline period (2011–2013)
- Similar baseline levels for priority outcomes
- Similar beneficiary characteristics on aggregate, such as mean age or Hierarchical Condition Category (HCC) score
- Similar health care markets, such as Health Resources and Services Administration (HRSA) scores measuring the degree of health professional primary care shortage in the PUMA or the degree of hospital market concentration within the PUMA[8]
- Similar characteristics—such as percentage of people living in multi-unit homes—that can make areas more vulnerable to disease outbreaks (we included these variables to mitigate risk of bias because of COVID-19; Appendix C provides details)
- Similar proportions of beneficiaries who are Black and who live in urban versus rural areas, and similar levels and trends for select outcomes for these subgroups of beneficiaries (this similarity should help make future estimates by beneficiary subgroups more credible)

In addition, we identified what we would count as sufficient balance for each of the matching variables. The method we used to reweight comparison PUMAs allowed us to set balance standards for each individual variable. We chose tight standards (< 0.15 standardized differences between the intervention and control groups) for trends in many baseline outcomes (because tight balance underlies the parallel trends assumptions) and for some variables needed for face validity or subgroups. We chose more relaxed standards (0.25 standardized differences or larger) for other types of variables, or in cases in which tight balance was not feasible without substantially affecting the quality of the comparison group in other ways—mainly reducing the size or geographic distribution of the group.

[8] We did not seek to match on participation in other alternative payment models (such as Accountable Care Organizations, Comprehensive Primary Care Plus, etc.) because most of these programs had not yet begun during our baseline period (2011–2013) or had low participation.

B.1.2c. Decision to include baseline outcomes in matching

We tried developing a comparison group without matching on baseline outcomes and trends to reduce the risk that regression-to-the mean could bias impact estimates.[9] The resulting comparison groups, however, had substantially different levels and trends than the intervention group, which creates its own risk of bias if such non-parallel trends persist into the intervention period. We chose to match on outcomes (levels and trends) to improve balance on these variables and because several other aspects of the design help to mitigate the risk of regression-to-the mean. First, we used PUMAs with a large number of Medicare beneficiaries, substantially limiting the noise that underlies regression-to-the mean bias. Second, we matched on outcomes over three years, rather than a single year, further limiting noise. Finally, we assessed whether the outcome means for the comparison group in 2010 (the year before the baseline period), moved away from the baseline trend line, as you would expect it would if regression-to-the mean were biasing the estimates. We did not see any evidence that the outcome in 2010 diverged substantially from the 2011–2013 trend.

We did not match on baseline levels or trends for hospital spending or total Medicare FFS spending, even though these are important outcomes in the evaluation. Hospital spending is difficult to match on because hospital spending in Maryland during the baseline period was much higher than in the rest of the country because of the all-payer rate-setting system. If we tried to match on hospital spending, the comparison group would likely be very small and have high hospital spending for reasons quite different than those in Maryland. As a result, the trends in hospital spending for such a comparison group likely would not reflect a reasonable counterfactual for Maryland. Similarly, total Medicare spending cannot be a priority matching variable because hospital spending accounts for more than half of total spending. We did include standardized hospital spending in our matching, which is calculated for Maryland and the comparison group by re-pricing claims to a standardized national fee schedule. In this way, standardized hospital spending is more closely related to hospital utilization than spending because the pricing effects have been removed.

[9] In difference-in-differences analyses, matching on outcomes can unintentionally create biased estimates if (1) there is random variation in outcome levels in the intervention and comparison units, and (2) the selected comparison units have a long-term mean that differs from the intervention group, but they are selected because they—randomly—look like the intervention group units at the time of matching. In these cases, the mean for the comparison group can snap back to its long-term mean in the post-intervention period, leading to post-intervention outcome differences that would be misinterpreted as model impacts (Daw and Hatfield 2018).

Exhibit B.1. Baseline measures for selecting PUMAs into the matched comparison group

Domain and measure	Data source	SD required[a]	Time period: 2013	Time period: Trend 2011–2013
Characteristics of Medicare beneficiaries in a PUMA				
Average age	Medicare enrollment	0.25	X	
Percentage Black	Medicare enrollment	0.15	X	
Percentage Black, 5 category distribution	Medicare enrollment	0.15	X	
Percentage non-Hispanic White	Medicare enrollment	0.25	X	
Percentage Hispanic	ACS	0.25	X	
Percentage female	Medicare enrollment	0.25	X	
Percentage with rural residence	Medicare enrollment U.S. Census	0.25	X	
Percentage with rural residence, 5-category distribution	Medicare enrollment U.S. Census	0.15	X	
Percentage with original reason for Medicare entitlement: disability, ESRD	Medicare enrollment	0.25	X	
Average HCC risk score	Medicare claims	0.25	X	
Percentage of FFS beneficiaries with diabetes	Medicare claims	0.25	X	
Characteristics of a PUMA and its population[b]				
Cost of living adjusted percentage below poverty level	ACS U.S. Census	0.25	X	
Percentage living in multi-unit structure, mobile home, or group quarters	ACS	0.33	X	
Percentage older than 64	ACS	0.33	X	
Percentage younger than 18	ACS	0.33	X	
Percentage speaks English well	ACS	0.33	X	
Percentage living in crowded home	ACS	0.33	X	
Percentage without a vehicle	ACS	0.33	X	
Percentage with high school degree (or equivalent)	ACS	0.33	X	
Percentage of all adults with diabetes	BRFSS	0.25	X	
Obesity prevalence, 2011-2013[c]	BRFSS	0.25	X	
Characteristics of a PUMA's health care system and insurance market				
Health professionals shortage area index score (for primary care providers)	HRSA	0.25	X	
Herfindahl-Hirschman Index (a measure of market concentration)	Medicare claims	0.25	X	
Percentage of residents in Medicare	Medicare enrollment ACS	Medium	X	
Number of primary care providers per 1,000 beneficiaries	Medicare claims	0.25	X	

Domain and measure	Data source	SD required[a]	Time period: 2013	Time period: Trend 2011–2013
Characteristics of hospitals in a PUMA				
PUMA has one or more acute care hospitals	Hospital Compare	0.25	X	
Percentage of all discharges from a major teaching hospital	Hospital Compare and Medicare claims	0.25	X	
Percentage of all discharges from hospitals belonging to a health care system	AHRQ CHSP and Medicare claims	0.25	X	
Total number of beds	Hospital Compare IPPS	0.25	X	
Characteristics of practices and providers in the PUMA				
Percentage of PCPs in practices (TINs) that are small (1 NPI in TIN)	Medicare claims MD-PPAS	0.25	X	
Percentage of PCPs in practices (TINs) that are large (6+ NPIs in TIN)	Medicare claims MD-PPAS	0.25	X	
Characteristics related to key subgroups living within the PUMA				
Average HCC score among Black beneficiaries	Medicare enrollment Medicare claims	0.25	X	
Average HCC score among non-Hispanic White beneficiaries	Medicare enrollment Medicare claims	0.25	X	
Outcomes related to Medicare FFS spending				
Non-hospital spending, 2013	Medicare claims	0.25	X	
Non-hospital spending, 2011–2013 trend	Medicare claims	0.15		X
Post-acute care spending, 2013	Medicare claims	0.25	X	
Post-acute care spending, 2011–2013 trend	Medicare claims	0.10		X
Outcomes related to health care utilization				
Standardized hospital spending, 2013	Medicare claims	0.25	X	
Standardized hospital spending, 2011–2013 trend	Medicare claims	0.15		X
Number of all-cause acute care hospital admissions, 2013	Medicare claims	0.25	X	
Number of all-cause acute care hospital admissions, 2011–2013 trend	Medicare claims	0.15		X
Number of outpatient emergency department visits and observation stays, 2013	Medicare claims	0.25	X	
Number of outpatient emergency department visits and observation stays, 2011–2013 trend	Medicare claims	0.15		X
Outcomes related to quality of care				
30-day post-discharge unplanned readmission, 2013	Medicare claims	0.25	X	
30-day post-discharge unplanned readmission, 2011–2013 trend	Medicare claims	0.15		X
Timely follow-up after a discharge for acute exacerbations of chronic conditions, 2013	Medicare claims	0.25	X	

Domain and measure	Data source	SD required[a]	Time period: 2013	Time period: Trend 2011–2013
Timely follow-up after a discharge for acute exacerbations of chronic conditions, 2011–2013 trend	Medicare claims	0.15		X
Number of potentially preventable admissions, 2013	Medicare claims	0.25	X	
Number of potentially preventable admissions, 2011–2013 trend	Medicare claims	0.10		X
Patients' rating of their hospital care, 2013	Medicare claims	0.50	X	
Patients' rating of their hospital care, 2011–2013 trend	Medicare claims	0.15		X
Beneficiaries' rating of their primary care physician, 2013	Medicare claims	0.25	X	

Note: We conducted matching at the region (PUMA) level. When applicable, we aggregated data to the PUMA level before analyzing or matching. For example, claims- and survey-based variables measured at the beneficiary or respondent level (respectively) in the underlying data files were aggregated to the PUMA-year level for matching. Hospitals' characteristics were aggregated accounting for hospital sizes.

[a] The column "SD" refers to the maximum standardized differences we allow between Maryland and the comparison group. In our reweighting algorithm we can set tolerances for individual variables to be more (lower SD) or less (higher SD) similar between Maryland and the control group (see the section on reweighting method below for more details). We aimed for a standard of 0.25 SDs where possible, but some variables were too difficult to match on (required large tradeoffs in balance elsewhere or size of the comparison group) and thus were allowed to be more dissimilar on standardized differences (e.g., patients' rating of their hospital care in 2013).

[b] To reduce the chance that statistical noise will affect survey-based and hospital-level measures, we used three-year averages rather than data from a single year.

[c] Obesity prevalence is the 2012 BRFSS files that used smoothed average from years 2011-2013.

ACS = American Community Survey; AHRQ CHSP = Agency for Healthcare Research and Quality Comparative Health System Performance (CHSP) Initiative; BRFSS = Behavioral Risk Factor Surveillance System; ESRD = end-stage renal disease; FFS = fee for service; HCC = Hierarchical Condition Category; HRSA = Health Resources and Services Administration; IPPS = Inpatient Prospective Payment System; MD-PPAS = Medicare Provider Practice and Specialty; PCP = primary care physician; PUMA = Public Use Microdata Area; SD = standardized differences; TIN = Taxpayer Identification Number.

B.1.3. Reweighting comparison PUMAs to create the matched comparison group

To select our comparison group, we used a method called stable balancing weights developed by Zubizarreta (2015). This method belongs to a class of methods called *minimal dispersion approximately balancing weights*, or *minimal weights* for short, that reweight the comparison group units without explicitly modeling the propensity to receive the intervention (that is, propensity scores).[10] Instead of modeling propensity scores, these methods find the weights that directly optimize certain attributes of the weights, targeting covariate balance directly and simultaneously minimizing a measure of dispersion of the weights. In the case of stable balancing weights, the optimization finds the weights for comparison units that achieve preset criteria for balance on individual matching variables while minimizing the dispersion of weights across the comparison units.

Reweighting and matching methods for constructing a comparison group are closely related conceptually (Stuart 2010). The methods have similar objectives and are based on similar principles. The main difference is that matching selects a subset of potential comparison regions

[10] Chattopadhyay et al. (2020) provided an overview of the *minimal weights* methods and contrasted them with more traditional inverse probability weighting approaches. Wang and Zubizarreta (2019) provided theoretical results.

to form the comparison group (and thus does not use all the available data), and weighting methods use all comparison regions and give different regions more or less weight (thus using all the available data, though some PUMAs can receive zero weight).[11]

The stable balancing weights method offered two main advantages over traditional matching techniques in the MD TCOC evaluation:

1. *It allows matching on the many variables identified in Exhibit B.1 as priorities.* Traditional matching methods based on propensity scores would likely not be able to match on so many variables because the propensity score model would risk overfitting with so many explanatory variables for only 44 intervention PUMAs.
2. *It allows for tailored balance criteria for each matching variable*. This tailoring enabled us to identify and make precise tradeoffs between balance on select variables versus the size and distribution of the comparison group.

Using an optimization-based approach, theoretically, any number of criteria can be set as constraints. As we add constraints (or tighten or require greater similarity between treated and comparison groups), however, the optimization problem becomes more difficult. The tradeoff to higher degrees of similarity across many different criteria is often the size of the comparison group represented. In other words, the algorithm will start to drop (that is, assign zero weight to) units that are too different from its target when there are no other options.

B.1.4. Assessing the quality of the matched comparison group in terms of size, geographic spread, balance, and statistical power

In selecting the comparison group, we aimed for a group that:

1. Was large and spread across much of the country, both to improve statistical power to detect effects and to avoid the possibility that idiosyncratic health shocks in any one area would drive the results.
2. Had sufficient balance on all variables listed as priorities for matching.
3. Had sufficient statistical power to detect policy-relevant impacts.

Conditions (1) and (2) generally trade off with one another—with more precise balance coming at the expense of a smaller and less geographically disperse comparison group. We explored several alternative comparison groups with CMS and selected the one that represented the best tradeoffs across these three dimensions.

[11] One way to think about it is that reweighting creates a matched comparison group (that is, a comparison group similar to the intervention group). Another is that matching is a form of reweighting (in which the weights for a region could be as simple as a 0 or 1).

B.1.4a. Size and geographic spread of the selected comparison group

The selected comparison group is large and covers much of the country. Exhibit B.2 shows several statistics that give a sense of the matched comparison group on a national scale. For example, the comparison group includes 37 states with a non-zero weight, and about two-thirds of the weight concentrates in the top 10 states (Exhibit B.3 shows weights of top 10 states). We also see that, in total, about 25% of the nation's Medicare FFS population has a positive weight in our comparison group (553 PUMAs), with 338 individual PUMAs accounting for about 90% of the total weight (Exhibit B.4 shows the weight of the top 10 PUMAs in our comparison group). Finally, we also display the effective sample ratio, which is an estimate of the ratio of comparison to treatment units that accounts for the sum of the weights and the dispersion of those weights (Exhibit B.2). Effective sample ratios of greater than 3:1 are generally considered to maximize the statistical power to detect effects for any given intervention group size. Higher ratios (for example, 10:1) only modestly increase statistical precision and can come at the cost of substantially worse balance on matching variables. Exhibit B.5 shows our final comparison group visually on a map of individual PUMAs. The more populous areas of the country have PUMAs with relatively small areas in the map, so populous PUMAs that received substantial weight might be hard to discern in this nationwide map.

Exhibit B.2. Matched comparison group diagnostics

Statistic	Value
Number of states with non-zero weight	37
Total weight of top 10 states	66.7%
Percentage of Medicare FFS population outside Maryland with a positive weight	24.9%
Total weight of top 50 PUMAs	29.4%
Number of PUMAs accounting for 50% of total weight	111
Number of PUMAs accounting for 90% of total weight	338
Effective sample ratio comparison: treatment	7.05

FFS = fee for service; PUMA = Public Use Microdata Area.

Exhibit B.3. Percentage of the selected comparison group in the top 10 most highly weighted states

State	PUMA FFS population (as % of total Medicare FFS population)	Comparison group FFS population (as % of Medicare FFS beneficiaries in the comparison group)
Illinois	5.0	12.4
Virginia	3.0	11.9
New Jersey	3.3	10.6
Georgia	2.9	6.3
New York	5.7	6.6
North Carolina	3.8	4.8
Pennsylvania	4.0	3.7
Florida	6.8	3.7
Connecticut	1.3	3.3
Texas	6.8	3.2

Note: As an example, 5% of the nation's Medicare FFS beneficiaries live in Illinois. In contrast, 12% of the Medicare FFS beneficiaries in the comparison group live in Illinois.

FFS = fee for service; PUMA = Public Use Microdata Area.

Exhibit B.4. Comparison group statistics and top 10 PUMAs by final analysis weight

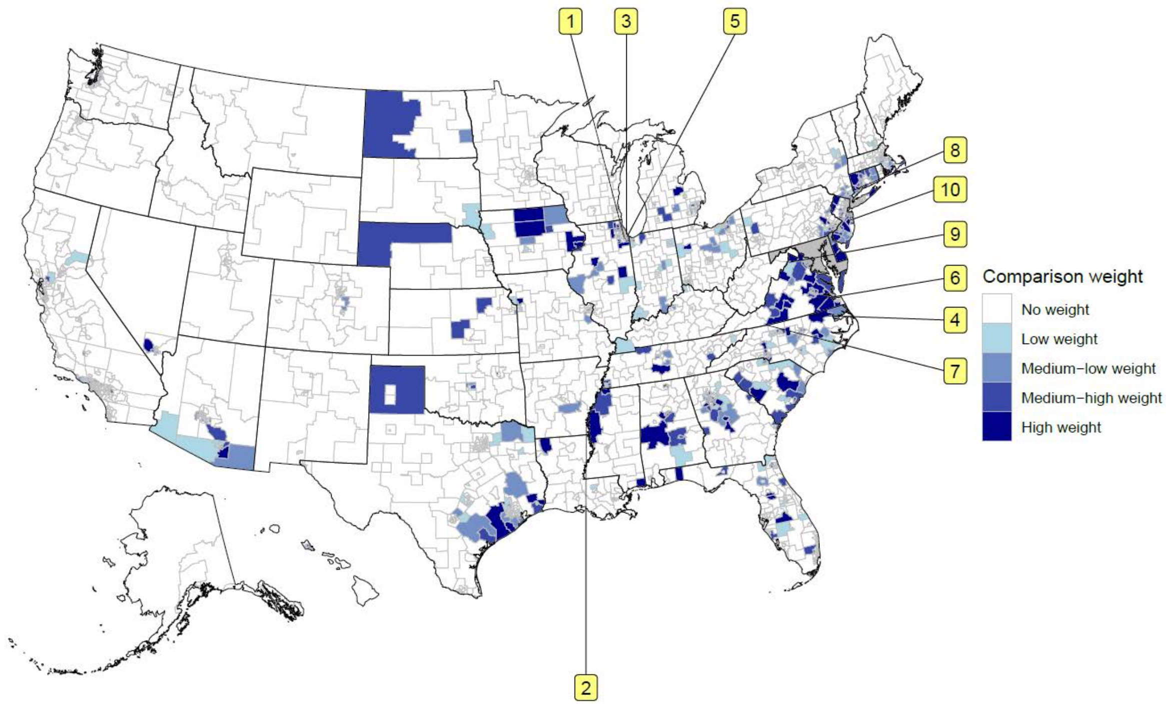
PUMA	PUMA FFS population[a]	Selected comparison group[b]
Illinois—Cook County (South) —Bloom and Rich Townships	21,426	1.26%
Mississippi—South Delta Region	19,821	0.93%
Illinois—Chicago City (South)--Auburn Gresham, Roseland, Chatham, Avalon Park and Burnside	16,842	0.88%
North Carolina—Halifax, Hertford, Northampton and Warren (East) Counties	20,711	0.87%
Illinois—Cook County (Southeast)--Thornton Township	18,735	0.86%
Virginia—Crater Planning District Commission	24,582	0.82%
Virginia—West Piedmont Planning District Commission	35,061	0.82%
Connecticut—Danbury, Ridgefield, Bethel, Brookfield, New Fairfield, Redding and Sherman Towns	18,560	0.79%
Virginia—Northern Shenandoah Valley Regional Commission (North)	16,017	0.78%
New Jersey—Ocean County (Northwest)	25,331	0.76%

[a] Number of Medicare FFS beneficiaries living in this PUMA in 2013.

[b] Final weight that the PUMA receives, which is a combination of the final matching weight and the PUMA size (FFS Medicare population).

FFS = fee for service; PUMA = Public Use Microdata Area.

Exhibit B.5. Map showing which PUMAs received positive weights in the selected comparison group



Note: Some PUMAs with very small areas are difficult to see on this map (for example, areas around Los Angeles, Chicago, and other major cities). As such, it might be difficult to see some PUMAs getting significant weight from this map alone. Yellow markers [1-10] indicate the top 10 PUMAs by weight in the comparison group. No weight = 0 weight; Low weight = weights in the 1st (0-25% quartile); Medium-low weight = weights in 2nd quartile; Medium-high weight – weights in the 3rd quartile; high weight – weights in the 4th quartile.

PUMA = Public Use Microdata Area.

B.1.4b. Balance on matching variables

Overall, we achieved good balance in our selected comparison group, including on outcomes, with most measures no more than 0.25 standardized differences apart from Maryland (Exhibit B.6).[12] The stable balancing weights method ensures all measures included in the algorithm meet the selected criteria (or the algorithm would fail). That is, we achieved balance that was *no worse* than the balance criteria specified in Exhibit B.6 and, in many cases, significantly better. In general, the balance criteria can be used to assess the relative importance we assigned to an individual variable in our matching algorithm. Smaller standardized differences represent tighter balance. We also included in Exhibit B.6 several variables we chose not to match on explicitly (for example, COVID-19-related variables) but that we were interested in checking balance on.

[12] Throughout our matching process, we intentionally calculated standard deviations used in constructing standardized differences at the PUMA level, rather than at the beneficiary level, as is often seen in final balance tables for beneficiary-level regressions. PUMA-level standard deviations are much smaller than beneficiary-level standard deviations, especially for measures such as HCC scores or beneficiary outcomes. This choice results in much stricter requirements on the standardized differences scale. We took this approach to be conservative, and because our comparison group is constructed at the PUMA level—a higher level of aggregation—we included several matching variables that are measured at the PUMA level. Because of these matching criteria, we achieved good balance for beneficiary-level measures in our final regressions.

Exhibit B.6. Balance between Maryland and selected comparison group on key characteristics and outcomes

Variable description	Maryland mean	National pre-weighted mean	Standardized difference pre-weighting	Balance criteria (standardized differences)	Selected comparison group (National post-weighted mean)	Standardized difference post-weighting
Percentage with rural residence	16.35	25.78	-0.36	0.25	15.15	0.05
Percentage with rural residence, category [0,1]	0.41	0.27	0.32	0.15	0.41	0.00
Percentage with rural residence, category [1,25]	0.25	0.30	-0.12	0.15	0.32	-0.15
Percentage with rural residence, category [25,50]	0.26	0.20	0.15	0.15	0.20	0.15
Percentage with rural residence, category [50,75]	0.08	0.18	-0.27	0.15	0.07	0.03
Percentage with rural residence, category [75,100]	0.00	0.05	-0.23	0.15	0.00	-0.01
Average age	71.47	70.58	0.43	0.25	71.23	0.11
Percentage female	57.46	55.49	0.87	0.25	56.89	0.25
Percentage non-Hispanic White	70.58	80.21	-0.49	0.25	71.29	-0.04
Percentage Black	22.74	9.66	0.88	0.15	20.52	0.15
Percentage Black, category [0, 5]	0.20	0.57	-0.74	0.15	0.28	-0.15
Percentage Black, category [5, 15]	0.36	0.24	0.29	0.15	0.34	0.04
Percentage Black, category [15, 25]	0.18	0.09	0.33	0.15	0.14	0.15
Percentage Black, category [25, 50]	0.08	0.08	-0.02	0.15	0.09	-0.04
Percentage Black, category [50, 100]	0.18	0.03	0.89	0.15	0.16	0.15
Percentage Hispanic	7.30	13.47	-0.39	0.25	11.23	-0.25
Percentage with original reason for Medicare entitlement: disability, ESRD	20.07	25.59	-0.61	0.25	22.33	-0.25
Average HCC risk score	1.11	1.11	0.03	0.25	1.13	-0.12
Percentage of adults with diabetes	9.33	9.19	0.08	0.25	9.19	0.08
Percentage with diabetes	30.12	28.33	0.32	0.25	29.54	0.10
Percentage below federal poverty level (adjusted for cost of living)	14.07	16.75	-0.36	0.25	15.92	-0.25
Percentage of residents in Medicare	15.03	17.85	-0.38	0.25	16.43	-0.19
PUMA has one or more acute care hospitals	0.71	0.87	-0.47	0.25	0.79	-0.25
Number of hospital beds	245.92	339.89	-0.27	0.25	333.92	-0.25

Variable description	Maryland mean	National pre-weighted mean	Standardized difference pre-weighting	Balance criteria (standardized differences)	Selected comparison group (National post-weighted mean)	Standardized difference post-weighting
Percentage of all discharges from a major teaching hospital	14.35	15.14	-0.04	0.25	19.36	-0.25
Herfindahl-Hirschman Index	0.29	0.25	0.30	0.25	0.27	0.18
Percentage of PCPs in practices (TINs) that are small (1 NPI in TIN)	28.11	21.97	0.37	0.25	26.35	0.11
Percentage of PCPs in practices (TINs) that are large (6+ NPIs in TIN)	46.84	49.74	-0.12	0.25	46.49	0.01
Number of primary care providers per 1,000 beneficiaries	6.14	5.05	0.22	0.25	5.06	0.21
Average health professionals shortage area index score (for primary care providers)	3.36	4.54	-0.22	0.25	3.62	-0.05
Percentage living in multi-unit structure, mobile home, or group quarters	19.04	25.60	-0.48	0.33	23.55	-0.33
Percentage older than age 64	13.86	14.97	-0.26	0.33	14.34	-0.11
Percentage younger than age 18	22.52	22.84	-0.09	0.33	23.21	-0.20
Percentage with high school degree (or equivalent)	92.06	89.65	0.44	0.33	91.16	0.17
Percentage speaks English well	97.57	96.43	0.25	0.33	96.88	0.15
Percentage living in crowded home	6.45	9.00	-0.47	0.33	8.04	-0.29
Percentage without a vehicle	7.03	6.34	0.09	0.33	7.36	-0.04
Non-hospital spending, 2013	5,299.52	5,185.46	0.09	0.25	5,609.06	-0.25
Non-hospital spending, 2011–2013 trend	-40.08	-71.05	0.30	0.15	-55.60	0.15
Medicare Part A post-acute care spending, 2013	1,115.93	1,075.02	0.12	0.25	1,160.69	-0.14
Medicare Part A post-acute care spending, 2011–2013 trend	9.23	8.66	0.01	0.10	12.65	-0.05
Standardized hospital spending, 2013	4,593.70	4,562.98	0.04	0.25	4,647.01	-0.07
Standardized hospital spending, 2011–2013 trend	98.87	69.36	0.28	0.15	83.05	0.15
Number of all-cause acute care hospital admissions, 2013	320.49	304.03	0.30	0.25	314.65	0.11
Number of all-cause acute care hospital admissions, 2011–2013 trend	-20.15	-14.71	-0.58	0.15	-18.74	-0.15
Number of outpatient ED visits and observation stays, 2013	461.94	498.73	-0.29	0.25	454.57	0.06

Variable description	Maryland mean	National pre-weighted mean	Standardized difference pre-weighting	Balance criteria (standardized differences)	Selected comparison group (National post-weighted mean)	Standardized difference post-weighting
Number of outpatient ED visits and observation stays, 2011–2013 trend	11.88	9.13	0.19	0.15	11.55	0.02
30-day post-discharge unplanned readmission, 2013	17.85	16.34	0.62	0.25	17.24	0.25
30-day post-discharge unplanned readmission, 2011–2013 trend	-0.76	-0.45	-0.50	0.15	-0.66	-0.15
Timely follow-up after a discharge for acute exacerbations of chronic conditions, 2013	67.36	67.92	-0.10	0.25	67.76	-0.07
Timely follow-up after a discharge for acute exacerbations of chronic conditions, 2011–2013 trend	0.80	0.79	0.00	0.15	1.00	-0.15
Number of potentially preventable admissions, 2013	56.96	54.88	0.13	0.25	56.04	0.06
Number of potentially preventable admissions, 2011–2013 trend	-4.22	-3.53	-0.23	0.10	-4.23	0.01
Patients' rating of their hospital care, 2013	66.51	70.41	-0.93	0.50	68.61	-0.50
Patients' rating of their hospital care, 2011–2013 trend	0.35	1.01	-0.64	0.15	0.50	-0.15
Beneficiaries' rating of their primary care physician, 2013	89.88	89.99	-0.04	0.25	90.23	-0.13
Obesity prevalence, 2012	28.24	27.95	0.06	0.25	27.03	0.25
Number of COVID-19 outpatient ED visits and observation stays, 2020	6.40	9.07	-0.51	Not included in matching	7.56	-0.22
Number of COVID-19 outpatient ED visits and observations stays, 2021	10.30	12.30	NA	Not included in matching	10.56	NA
Excess number of all-cause acute care hospital admissions (2020 minus 2019)	-39.69	-39.95	0.02	Not included in matching	-40.49	0.06
Excess number of outpatient ED visits and observation stays (2020 minus 2019)	-121.10	-114.80	-0.19	Not included in matching	-117.19	-0.12
Excess number of ED visits and observation stays ending in an inpatient stay (2020 minus 2019)	-27.08	-24.16	-0.23	Not included in matching	-25.36	-0.14
Number of surgical hospitalizations (2020 minus 2019)	-14.75	-14.10	-0.14	Not included in matching	-14.22	-0.12
Number of elective hospitalizations (2020 minus 2019)	-11.46	-11.73	0.05	Not included in matching	-11.31	-0.03

Variable description	Maryland mean	National pre-weighted mean	Standardized difference pre-weighting	Balance criteria (standardized differences)	Selected comparison group (National post-weighted mean)	Standardized difference post-weighting
Average HCC score, Black beneficiaries	1.34	1.42	-0.52	0.33	1.39	-0.33
Average HCC score, non-Hispanic White beneficiaries	1.05	1.06	-0.09	0.33	1.06	-0.04
Number of potentially preventable admissions, Black beneficiaries, 2013	80.05	77.54	0.13	Not included in matching	76.92	0.16
Number of potentially preventable admissions, 2011–2013 trend, Black beneficiaries	-4.63	-4.63	0.00	Not included in matching	-5.51	0.11
Number of potentially preventable admissions, non-Hispanic White beneficiaries, 2013	51.90	52.99	-0.07	Not included in matching	51.64	0.02
Number of potentially preventable admissions, 2011–2013 trend, non-Hispanic White beneficiaries	-4.27	-3.41	-0.28	Not included in matching	-3.85	-0.14
30-day post-discharge unplanned readmission, Black beneficiaries, 2013	21.87	21.48	0.12	Not included in matching	21.98	-0.03
30-day post-discharge unplanned readmission, 2011–2013 trend, Black beneficiaries	-0.89	-0.49	-0.22	Not included in matching	-0.66	-0.13
30-day post-discharge unplanned readmission, non-Hispanic White beneficiaries, 2013	16.71	15.77	0.46	Not included in matching	15.97	0.36
30-day post-discharge unplanned readmission, 2011–2013 trend, non-Hispanic White beneficiaries	-0.80	-0.45	-0.53	Not included in matching	-0.66	-0.21

Notes: The pre-weighted means are the raw PUMA-level means (weighted only for FFS beneficiary count). Post-weighted means are weighted by the final matching weights. Standardized differences are a measured using the PUMA-level standard deviations.

ED = emergency department; ESRD = end-stage renal disease; FFS = fee for service; HCC = Hierarchical Condition Category; NPI = National Provider Identifier; PCP = primary care physician; PUMA = Public Use Microdata Area; TIN = Taxpayer Identification Number.

B.1.4c. Statistical power

Based on the size of the selected comparison group, and assumptions about the variation in outcomes (and degree of clustering within PUMAs), we estimated that the evaluation would have sufficient statistical power. For example, we estimated that the model would be able to reliably detect an impact on hospital admissions of 2.5% or larger.[13] The strong statistical power stems, in part, from the large size of the comparison group, as indicated by the effective sample size ratio of comparison to intervention group beneficiaries of 7:1.

The impact estimates shown in Chapter 2 confirm that the estimates have good statistical power, with the model finding impacts on total spending as small as 1.0% being statistically different from zero.

B.1.5. Unadjusted mean outcomes over time, for Maryland and the comparison group

To help interpret what drives the difference-in-differences impact estimates, we include the size of the intervention and comparison groups over time (Exhibit B.7) and the trends in unadjusted (but comparison group weighted) means for study outcomes since 2011 (Exhibits B.8-B.11) for these populations. The figures are especially helpful for identifying the time trends in the intervention and comparison groups that underlie the differences-in-differences impact estimates—for example, that all-cause admissions have been falling steadily in the intervention and comparison groups but more so in the intervention group (Exhibit B.10, Panel A). For most outcomes, the trends extend through 2022. For standardized spending, total spending including non-claims payments, and patients' ratings of their hospitals, the trends run through 2021 instead because of lags in data availability. For patients' rating of their personal doctor, trends extend through 2022 but omit 2020 because the survey was not collected in 2020 due to the COVID-19 pandemic. For the means figures, beneficiaries in Maryland are weighted by their observability in the year, and beneficiaries in the comparison group are weighted by their observability and their matching weights (see B.1.3). Because episodes-level outcomes (30-day unplanned readmission and timely follow-up after acute exacerbation of a chronic condition) are not annualized, episodes in Maryland receive a weight of 1, and episodes in the comparison group are weighted by their matching weights.

For the beneficiary and episode analyses, the ratio of comparison group to Maryland beneficiaries decreases slowly by year. This pattern occurs largely because more Medicare FFS beneficiaries enter Medicare Advantage (and exit the study population) over time in the comparison group than in Maryland, where rates of Medicare Advantage enrollment are low.[14] In contrast, the ratio for the patients' ratings of their personal doctor stays relatively constant over time because that analysis includes both FFS and Medicare Advantage beneficiaries.

[13] By reliably detect, we mean that the regressions would have 80% power to detect a difference of at least the size indicated (using a two-tailed test and a $p < 0.10$ cutoff for statistical significance).

[14] Section A.6 in the quantitative-only report from the model's first three years, located here, provides more detail on how differential enrollment into Medicare Advantage might bias our impact estimates, and the methods we used to limit that bias risk.

The unit of analysis for patients' rating of their hospital is the hospital—not the beneficiary—which is why the study population is so much smaller for that outcome. To be included in the analysis, the hospitals also needed to meet several inclusion criteria (see B.2.8c). For example, they needed to have Hospital Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey results in at least one year. CMS only reports Hospital CAHPS data for hospitals with inpatient beds (so free-standing emergency rooms could not be included) and meet a minimum threshold for number of respondents. These filters are the reason the number of hospitals in the study population for Maryland for this analysis is about 40, smaller than the 52 hospitals in the Maryland state agreement.[15]

Exhibit B.7. Size of the Maryland and comparison groups over time (weighted)

Year	Weighted Maryland count	Weighted comparison count	Comparison: Maryland weighted ratio
Beneficiary-level analysis counts (number of Medicare FFS beneficiaries)[a]			
2011	626,217	7,504,397	12.0
2012	645,971	7,614,399	11.8
2013	669,371	7,715,748	11.5
2014	691,471	7,743,808	11.2
2015	714,401	7,777,619	10.9
2016	723,334	7,900,042	10.9
2017	726,943	7,818,014	10.8
2018	735,084	7,778,738	10.6
2019	751,129	7,651,834	10.2
2020	759,095	7,480,686	9.9
2021	737,204	7,168,285	9.7
2022	721,580	6,906,045	9.6
Episode analysis: 30-day post-discharge unplanned readmission index admission counts (number of index admissions)			
2011	209,338	2,473,462	11.8
2012	202,245	2,378,216	11.8
2013	198,208	2,262,289	11.4
2014	194,090	2,184,900	11.3
2015	193,968	2,181,226	11.2
2016	190,312	2,162,173	11.4
2017	186,095	2,138,469	11.5
2018	181,046	2,095,779	11.6
2019	177,061	2,055,461	11.6
2020	147,451	1,687,878	11.4
2021	148,052	1,613,028	10.9
2022	144,949	1,559,944	10.8

[15] The one exception was a data anomaly in 2016 when several Maryland hospitals did not report Hospital CAHPS data.

Year	Weighted Maryland count	Weighted comparison count	Comparison: Maryland weighted ratio
Episode analysis: Timely follow-up after acute exacerbation of chronic conditions denominator counts (number of eligible index admissions and ED visits)			
2011	72,418	847,914	11.7
2012	74,493	848,486	11.4
2013	74,250	821,043	11.1
2014	75,463	812,035	10.8
2015	77,771	814,155	10.5
2016	72,246	776,766	10.8
2017	74,675	784,598	10.5
2018	73,071	755,780	10.3
2019	74,064	732,239	9.9
2020	55,670	527,239	9.5
2021	53,528	498,315	9.3
2022	50,598	465,247	9.2
Beneficiary perspective analysis: Patients' rating of their personal doctor counts (weighted number of surveys)[b]			
2011	538,795	7,421,178	13.8
2012	578,945	7,705,830	13.3
2013	607,638	7,902,335	13.0
2014	672,171	9,083,900	13.5
2015	669,598	8,605,234	12.9
2016	677,566	8,787,272	13.0
2017	635,609	8,059,773	12.7
2018	706,400	9,487,646	13.4
2019	648,009	8,904,294	13.7
2020	NA	NA	NA
2021	649,025	9,086,619	14.0
2022	675,091	9,449,682	14.0
Beneficiary perspective analysis: Patients' rating of their hospital counts (weighted number of hospitals)[c]			
2011	39	666	17.2
2012	39	668	17.2
2013	40	667	16.8
2014	41	669	16.2
2015	42	671	16.1
2016[d]	excluded	excluded	excluded
2017	42	674	16.2
2018	42	670	16.1
2019	41	660	16.0
2020	41	653	15.9
2021	41	650	15.8

[a] Beneficiary weighted counts are different from counts reported in Chapter 2 because they are weighted by observability (a beneficiary is given a weight of 0.5 if only observable in claims for half the year) here, and unweighted in the table in Chapter 2.

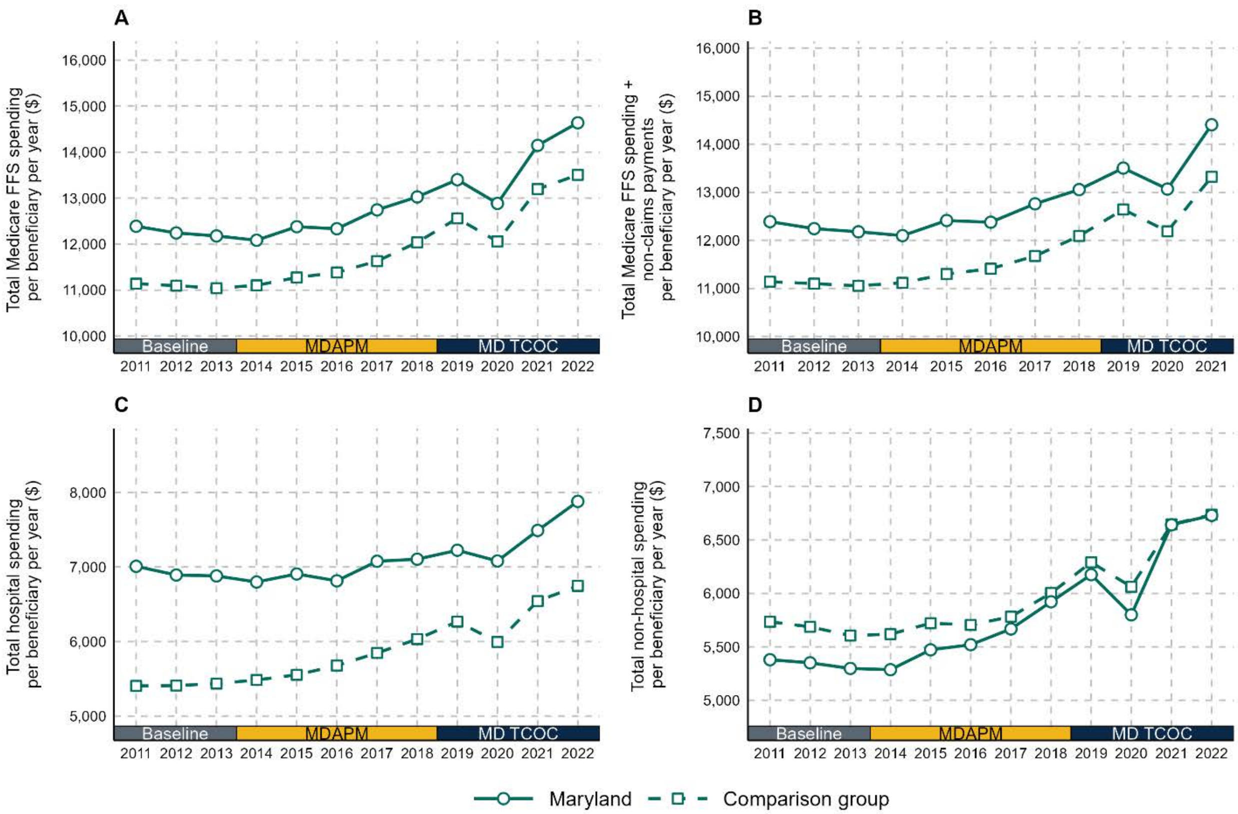
[b] The counts for patients' rating of their personal doctor are weighted by our PUMA matching weight (normalized to mean 1) multiplied by the CAHPS survey weights. CAHPS weights are designed to inflate back to approximately the total number of FFS and Medicare Advantage beneficiaries in the state, not the actual number of surveys completed. Numbers are lower than the total number of FFS beneficiaries in our beneficiary-level sample because not all people who take the survey respond to this question – only those with a primary doctor and who have received care in the last six months. Actual survey response rates are declining during this period from about 50% in 2011 to less than 33% among FFS beneficiaries in 2019.

[c] The counts for patients' rating of their hospital are weighted by our PUMA matching weight (normalized to mean 1) multiplied by the size of the hospital based on the number of discharges in 2013 or the year after the first year the hospital appears in our data (normalized to mean 1).

[d] Calendar year 2016 was excluded from analyses because several large hospitals in Maryland did not report scores in that year, potentially skewing results.

CAHPS = Consumer Assessment of Healthcare Providers and Systems; ED = emergency department; FFS = fee for service; PUMA = Public Use Microdata Areas. NA=not available.

Exhibit B.8. Unadjusted spending per beneficiary per year after matching

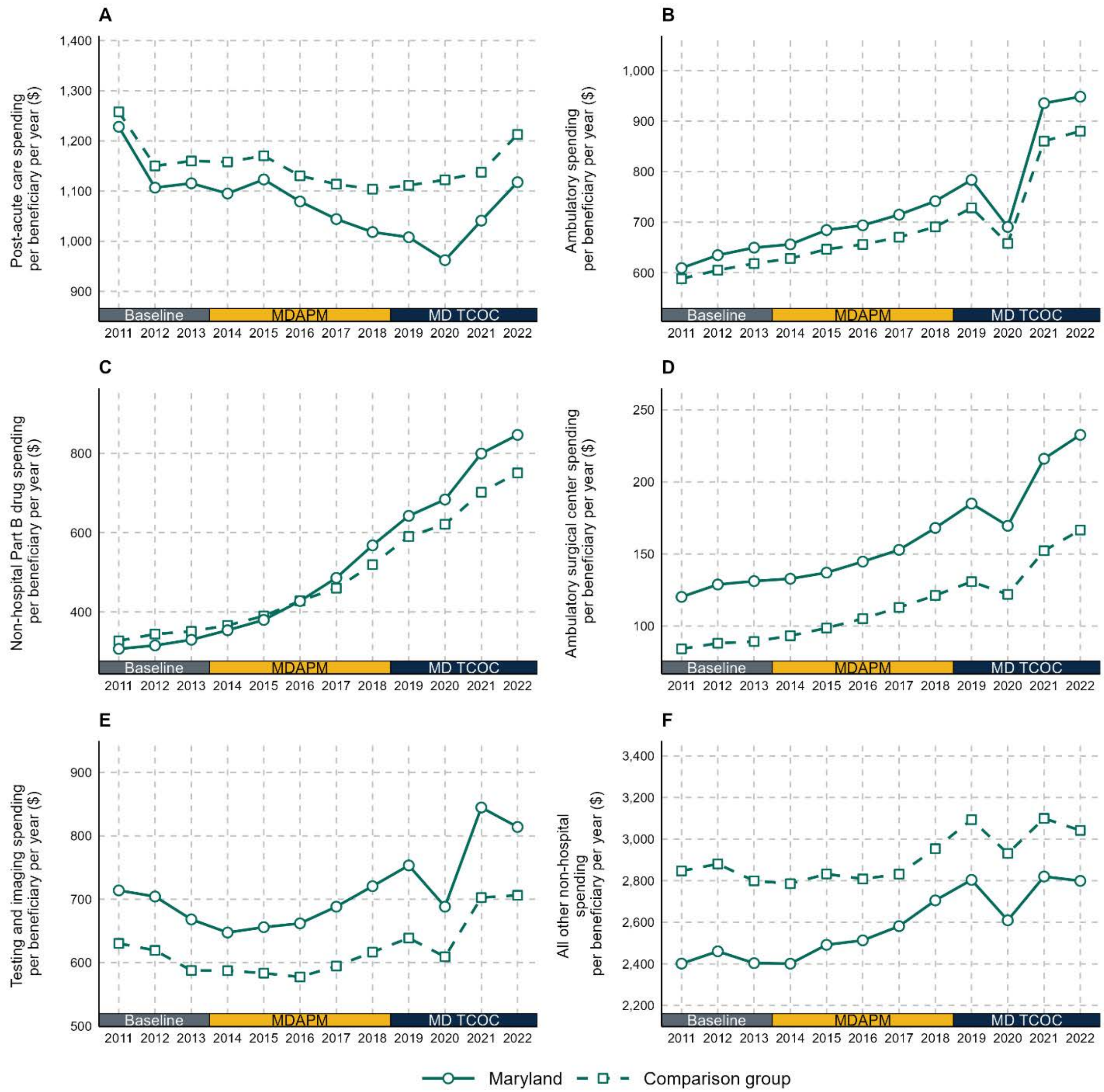


Source: Mathematica's analysis of Medicare FFS claims

Note: Maryland mean is weighted for observability in Medicare FFS claims. Comparison group mean is weighted for matching and observability.

FFS = fee for service.

Exhibit B.9. Unadjusted non-hospital spending per beneficiary after matching

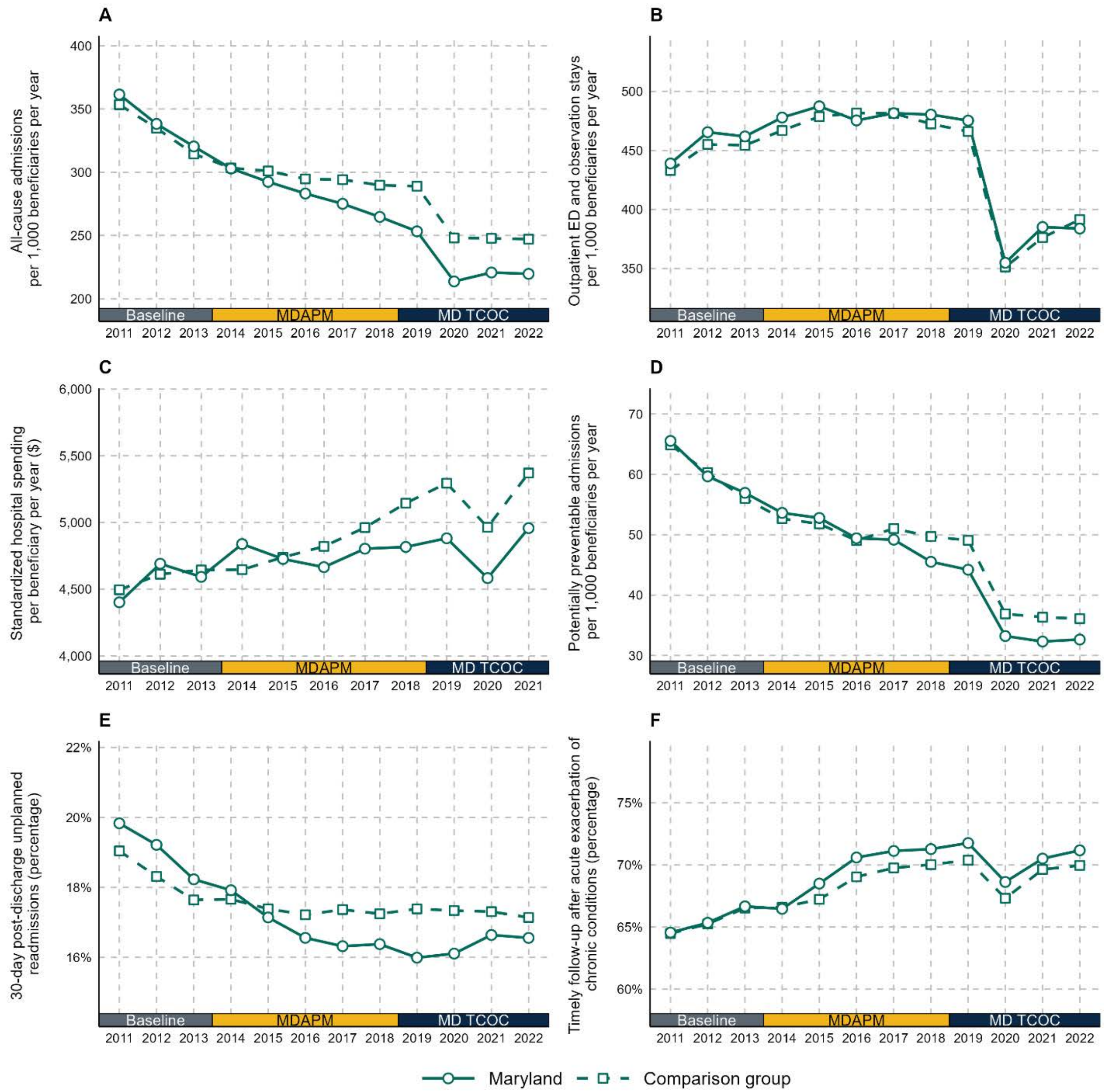


Source: Mathematica's analysis of Medicare FFS claims

Note: Maryland mean is weighted for observability in Medicare FFS claims. Comparison group mean is weighted for matching and observability.

FFS = fee for service.

Exhibit B.10. Unadjusted utilization and quality outcomes after matching

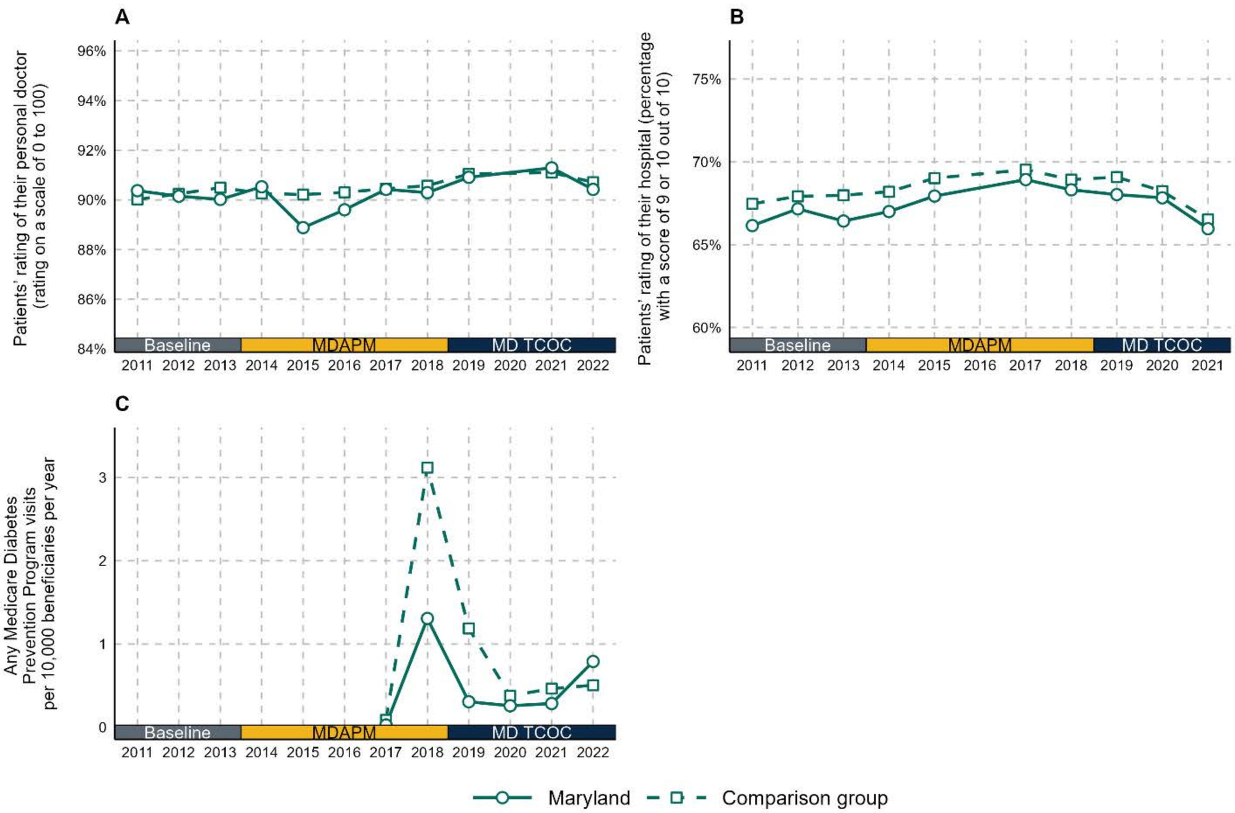


Source: Mathematica's analysis of Medicare FFS claims

Note: Maryland mean is weighted for observability (except for 30-day unplanned readmissions and follow-up after acute exacerbation which are episode level). Comparison group mean is weighted for matching and observability.

ED = emergency department; FFS = fee for service.

Exhibit B.11. Unadjusted patient experience and diabetes prevention program outcomes after matching



Source: Mathematica's analysis of Medicare FFS claims, FFS and Medicare Advantage patient-level CAHPS surveys, and hospital-level CAHPS surveys.

Note: For Diabetes Prevention Program Services, Maryland mean is weighted for observability. Comparison group mean is weighted for matching and observability. For patients' rating of their personal doctor, both Maryland and comparison groups are weighted by Medicare FFS CAHPS survey weights and the comparison group is also weighted for matching. For patients' rating of their hospital both Maryland and comparison hospitals are weighted by the number of Medicare FFS hospital discharges in 2013 and comparison group hospitals are also weighted for matching. For hospital rating, 2016 was excluded from analyses because several large hospitals in Maryland did not report scores in that year, potentially skewing results. Diabetes Prevention Program Services were not billable to Medicare until 2017.

CAHPS = Consumer Assessment of Healthcare Providers and Systems; FFS = fee for service.

B.1.6 Regression model specifications

B.1.6a. Regression specifications and statistical testing for beneficiary-year and episode-year Medicare FFS claims-based analyses

We used linear regression models to implement the difference-in-differences impact analyses. We measured impacts separately for each year and separately for the MDAPM and MD TCOC periods. The findings in this report included three units of analysis: (1) analyses of observations for each Medicare FFS beneficiary in Maryland and the matched comparison regions for each year (beneficiary-year analyses, including patients' rating of their personal doctor), (2) analyses of episode outcomes with observations for each episode for each year (episode-year analyses), and (3) analyses of hospital ratings with observations for each hospital for each year they appear in the data. The beneficiary-year and episode-year models accounted for the clustering of beneficiaries within PUMAs through cluster-robust standard errors, controlled for time-invariant effects of unobserved confounders and common shocks through the use of fixed effects, and they included baseline and time-varying covariates as independent variables.

Impact estimates

The difference-in-differences regression models for the beneficiary-year analyses with claims-based outcome measures used Medicare FFS data with one observation per beneficiary for each year (2011 to 2022). The regression models for the episode-year analysis took the same form, but with the unit of analysis as the episode rather than the beneficiary. The regression model to estimate the yearly impact for beneficiary- and episode-level estimates took the following form:

$$(1) \qquad y_{it} = \sum\nolimits_{\tau=2014}^{2022} T_\tau M_r \delta_{t,\tau} + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

In this model, y_{it} represents the outcome for beneficiary *i* (or episode *i*) in year *t* in region (PUMA) r, τ indexes years (with $\tau = 2011$ corresponding to the first year),[16] M_r equals 1 for Maryland beneficiaries (or episodes) and 0 for beneficiaries (or episodes) from the comparison regions, and T_τ is a dummy variable that equals 1 for observations in year τ and equals zero otherwise. X_{it} is a set of independent covariates whose relationship with the outcome we allow to change with time using an interaction term. The covariates are available in Table A.7. γ_t represents a set of year fixed effects and μ_r represents a set of PUMA-level fixed effects for beneficiary-year outcomes and hospital fixed effects for episode-year outcomes.

Beneficiaries in Maryland generally receive a weight of 1 in the regression models. But in cases in which a beneficiary is unobservable (that is, not alive or enrolled in Medicare Part A and B with Medicare as their primary payer) the whole year, we annualized their beneficiary-year outcomes and constructed observability weights that reflect the amount of time that the beneficiary is observable in the year. For the comparison group beneficiaries, we applied the matching weights (detailed in Sections B.1.2 and B.1.3) to account for the PUMA-level reweighting along with the observability weights; the two weights were multiplied together to produce a final, beneficiary-level weight. For episode analyses, we applied the matching weights to comparison group beneficiaries, and Maryland beneficiaries received a weight of 1 because episode analyses were not annualized.

The impact estimates are the δ's—the change in mean outcomes in the intervention group each year after accounting for the changes in the comparison group in the respective year (the γ_t's).

[16] All time trends are relative to the last year of the baseline period (2013), which is the reference year in the regression models.

Separate estimates for each year (that is, one δ per year) allowed for nonlinearity in the effects (for example, effects might not occur immediately or could level off or decline over time).

In addition to the yearly impact estimates, we also estimated the combined effect during the MD TCOC period. The regression model to estimate the combined 2019–2022 impact estimates took the following form:

$$(2) \qquad y_{it} = \sum_{\tau=2014}^{2022} T_\tau M_r \delta_{t,\tau} + T_{2019-2022} M_r \delta_\gamma + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

In this model, y_{it} represents the outcome for beneficiary *i* in year *t* in region (PUMA) r, τ indexes years (with $\tau = 2011$ corresponding to the first year), M_r equals 1 for Maryland beneficiaries and 0 for beneficiaries from the comparison regions, T_τ is a dummy variable that equals 1 for observations in year τ and equals zero otherwise, and $T_{2019-2022}$ is a dummy variable that equals 1 for observations in years 2019 to 2022. X_{it} is a set of independent covariates whose relationship with the outcome we allow to change with time using an interaction term. The covariates are listed below in Table A.7. γ_t represents a set of year fixed effects and μ_r represents a set of PUMA-level fixed effects for beneficiary-year outcomes and hospital fixed effects for episode-year outcomes. δ_γ represents the impact estimates during the MD TCOC period.

Finally, we also estimated models that produce an estimate of the *difference* between impacts during the MD TCOC period and the end of the MDAPM period. Models that estimate this difference took one of the two following forms:

$$(3) \qquad y_{it} = \sum_{\tau=2014}^{2016} T_\tau M_r \delta_{t,\tau} + T_{2017-2022} M_r \omega_\gamma + \sum_{\tau=2019}^{2022} T_\tau M_r \delta_{t,\tau} + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

$$(4) \qquad y_{it} = \sum_{\tau=2014}^{2016} T_\tau M_r \delta_{t,\tau} + T_{2017-2022} M_r \omega_\gamma + T_{2019-2022} M_r \delta_\gamma + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

Model (3) estimates the difference, between impacts during the MD TCOC period and the end of the MDAPM period *for each year of the MD TCOC period*. Model (4) estimates this same difference, as an average across the full MD TCOC period to date. The key new term in models (3) and (4) is $T_{2017-2022} M_r$. This term is 1 for Maryland observations in any year from 2017 to 2022 and 0 otherwise. Adding this term and including terms during the MD TCOC period alone, allows us to interpret the δ_t impact estimates during the MD TCOC period as *net* of effects during the last two years of MDAPM (2017-2018). We estimate models this way (instead of simply combining and subtracting estimates from models (1) or (2) above to generate the difference) to accurately generate confidence intervals (CIs) and *p*-values for the difference in effects between the MD TCOC period (or individual years in the MD TCOC period) and the effects at the end of MDAPM (2017-2018). All other terms in models (3) and (4) that are shared with models (1) and (2) are interpreted the same.

Covariates

The covariates in Equations 1 to 4 are included to account for trends in the intervention and comparison groups, improve the precision of the impact estimates, and net out effects of any observed residual differences in characteristics between the intervention and comparison groups. A full list of the covariates included in the claims-based beneficiary-year analyses and for the episode analyses for readmission and timely follow-up is available in Exhibit B.12. We control for beneficiaries' demographic characteristics (age, race, ethnicity, and sex) and Medicare enrollment characteristics (original reason for entitlement, and whether a new Medicare beneficiary in each year) and a measure of the beneficiary's PUMA Social Vulnerability Index (SVI) ranking in the regression models. By incorporating beneficiary characteristics X_{it} from claims and other data sources (including the characteristics of the region in which the beneficiary lives), we control for shifts in beneficiary characteristics over time unrelated to the model that, if unaccounted for, might lead to spurious conclusions. The vector of coefficients, β , control for these types of effects. Each of the characteristics in X_{it} are interacted with year to allow their relationship with the outcome to vary over time.

Some of the beneficiary characteristics we included in the list of covariates were indicators of a beneficiary's health status each year. We identified health status based on the presence (or absence) of 36 condition categories and 1 indicator for greater or equal to three Chronic Conditions Data Warehouse (CCW) conditions. We developed this list of 36 conditions as those that (1) were included in CMS CCW 27 chronic conditions active from 2011–2020 or 40 Other Chronic Health, Mental Health, and Potentially Disabling Conditions; (2) had a prevalence large enough to reliably estimate its association with outcomes in individual years (3) were not conceptually endogenous, (that is they were not conditions that the Maryland Model explicitly aims to reduce).[17,18]

[17] We excluded sickle cell disease, pressure ulcers and chronic ulcers, spinal cord injury, spina bifida and other congenital anomalies of the nervous system, muscular dystrophy, traumatic brain injury and nonpsychotic mental disorders due to brain damage, cerebral palsy, and learning disabilities due to very small prevalence in FFS Medicare claims (< 0.05%)

[18] Endogenous conditions are those whose prevalence might be changed by the Maryland Model. If we adjusted for changes in these conditions over time, we might adjust away impacts of the Maryland Model. We flagged CCW conditions related to diabetes and behavioral health as endogenous, particularly because of the focus on reducing body mass index and drug overdose deaths under MD TCOC (Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation 2020).

For the unplanned readmissions, we controlled for the index admission category and the beneficiary's health and chronic conditions covariates used in the beneficiary-level regressions.[19] For the timely follow-up outcome after acute exacerbations of chronic conditions, we controlled for the specific chronic conditions used to define the measure and the beneficiary's health and chronic conditions covariates used in the beneficiary-level regressions.

The regional fixed effects, μ_r, in beneficiary-year models net out the effects of any time-invariant differences between the regions in Maryland and the comparison regions.[20] Controlling for PUMA fixed effects implicitly accounts for all PUMA-level baseline measures we used in constructing the matched comparison group, including characteristics of the Medicare beneficiaries in the region, the region and its population, the region's health care system and insurance market, hospitals in the region, practices and providers in the region, and primary care providers in the region. Therefore, we do not include any additional PUMA-level variables as control variables in the regressions. Hospital fixed effects for the episode analyses account for all time-invariant differences between Maryland and the comparison hospitals (including the types of services they provide) and changes in hospitals' market shares over time. Collectively, these terms improve the precision of the impact estimates (the δ's) by reducing the amount of unexplained variation in the outcome (ε_{it}).

[19] The index admission categories include surgical or cardio respiratory or cardiovascular or neurology or medicine based on the procedure codes and principal diagnosis, per CMS/Yale technical specifications.

[20] The size of the data from our analytic files (nearly 100 million observations when stacked across years) means we must use SAS statistical analysis software to implement regressions on the Virtual Research Data Center. SAS has limited options for absorbing many dummy variables in the regression (such as the > 2,300 PUMA fixed effects). For computational feasibility, we run regressions by "de-meaning" the outcomes and all covariates at the PUMA level. That is, for each variable in an observation (including the outcome and all covariates), we replace the variable's value with the observed value minus the PUMA-specific mean (across all years) for that variable. This method is mathematically equivalent (in linear models) to adding PUMA fixed effects but considerably faster because it does not need to estimate the PUMA fixed effects explicitly in the regressions.

Exhibit B.12. Covariates for the impact analyses, by type of regression model or outcome

Domain and measure	Claims-based beneficiary-year-level outcomes	Readmission rates (episode analyses)	Timely follow-up (episode analyses)	Patients' rating of their doctor	Patients' rating of their hospital
Age	X	X	X	X	
Gender	X	X	X	X	
Race and ethnicity	X	X	X	X	
Social Vulnerability Index	X	X	X	X	
Original reason for Medicare entitlement[a]	X	X	X	X	
New Medicare beneficiary	X	X	X	X	
Rural residence	X	X	X	X	
Region (PUMA) in which the beneficiary resides	X			X	
Health conditions					
Health condition flags (based on CCWs)[b]	X		X		
Has three or more CCW conditions	X		X		
Self-reported health				X	
Hospital		X	X		
5 clinical cohorts for unplanned readmission measure case mix risk adjustment[c]		X			
6 chronic conditions used to define the timely follow-up after acute exacerbations of chronic conditions measure[d]			X		
Case mix index					X

[a] ESRD is measured in Medicare enrollment date in addition to claims. The ESRD category includes all beneficiaries with ESRD, and the Disability insurance benefits category does not.

[b] See text in section 2.2.2. for a full list of CCW conditions included

[c] Cohorts are surgical, cardiorespiratory, cardiovascular, neurology, or medicine. See the specifications for 30-day unplanned readmission developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2020) for details on index admission cohort assignment.

[d] The categories are asthma or hypertension or coronary artery disease or heart failure or chronic obstructive pulmonary disease or diabetes. See the IMPAQ Health (2018) specifications for more details on the chronic condition category assignment.

ESRD = end-stage renal disease; FFS = fee for service; CCW = Chronic Conditions Data Warehouse; PUMA = Public Use Microdata Area.

B.1.6b. Regression specifications and statistical testing for beneficiary-year and hospital-year patients' ratings analyses

In addition to the claims-based outcomes described above, we estimated impacts for two measures of patient experience from national surveys; one on patients' ratings of their personal doctor and one on patients' rating of their hospital care.

Patients' rating of their personal doctor

Patients' rating of their personal doctor comes from the FFS and Medicare Advantage CAHPS surveys administered by CMS. These data contain a beneficiary ID that links directly to Medicare claims (see Appendix B.2.1g for details). As such, the regression models estimating difference-in-differences impacts were similar to the beneficiary-year regression models described in Section B.1.6a.

Although similar, the regression models for this outcome differ from the models described in B.1.6a. in the following ways:

- **Time period.** The COVID-19 pandemic caused a suspension of the surveys in 2020, and resumed in 2021 and 2022. As such, we only have three years of estimates during the MD TCOC period, excluding 2020.
- **Weighting.** In our main regression models, Maryland beneficiaries are weighted by their observability weights, and comparison group beneficiaries are weighted by their observability weight times the matching weight (see Section B.1.3). For patients' rating of their personal doctor, we weighted Maryland beneficiaries by the CAHPS survey weight, which is designed to correct for survey response bias and returns the weighted population counts to approximate the total FFS and Medicare Advantage population in the state (see Section B.1.5 for counts).[21] We weighted the comparison beneficiaries by the product of the CAHPS survey weight and the matching weight. Observability weight are not applicable in the survey analysis because respondents cannot be partially observed.
- **Covariates.** In the analysis for patients' rating of their personal doctors, we included beneficiaries' demographic characteristics, Medicare enrollment characteristics, and a measure of the beneficiary's PUMA SVI ranking in the same way we did for other beneficiary- and episode-level outcomes. We also included a measure of the beneficiary's self-reported education because this information was available as part of the CAHPS surveys. We did *not* include, however, time-varying health condition controls as measured by CCW chronic conditions. The reason is that the sample of beneficiaries in this analysis includes FFS and Medicare Advantage beneficiaries. The primary reason for including time-varying health condition controls in our FFS sample was to correct for differential changes in the population because of beneficiaries leaving for Medicare Advantage (see Section B.1.8). Because we do not have beneficiaries leaving our sample when they switch to Medicare Advantage, we do not need to correct for health status differences that are attributable to the changing sample, so we exclude health condition controls from our regressions. Another reason we do not include these controls is that they are missing for all Medicare Advantage beneficiaries because the CCW conditions require at least of year of FFS claims lookback to identify the conditions using diagnosis codes. We do, however, include a measure of self-reported health in our regressions, available for all beneficiaries who completed the survey. We include self-reported health to help further correct for survey response bias because the correction made by the CAHPS weights is done at the state level, and our matched comparison group is defined at the PUMA level. Importantly, our results do not materially change with and without the inclusion of self-reported health as a covariate.

[21] The study population for the doctor rating outcome in Maryland is smaller than the total Medicare population (FFS and Medicare Advantage) in the state. This difference occurs largely because the survey only asks beneficiaries to rate their personal doctor if they say that they have a personal doctor who they have seen in the past six months.

Patients' rating of their hospital

Patients' rating of their hospital care comes from the Hospital CAHPS survey, which is administered by individual hospitals (or third-party contractors) to randomly selected patients recently discharged from the hospital (regardless of payer). The data from Hospital CAHPS is publicly available from CMS's website and stored at the hospital-year level (that is, an average set of responses for that hospital in the reporting period). The hospital rating measure that we used is defined as the percentage of survey respondents who rated their hospital overall 9 or 10 out of 10 (see Section B.2.1g for more details on the measure). Though the core of the difference-in-differences model we use is similar to the equations in Section B.1.6a, the regression specifications for patients' rating of their hospital care has several important differences.

First, because data are available only at the hospital-year level, we estimate regressions using hospital-year observations. Because hospitals are different sizes and might contribute differently to our estimate of impacts, this has implications for how we weight observations in our sample. We continue to use our PUMA-level matching weight to ensure we use the same comparison group as for other outcomes, but we then multiply the matching weight by the normalized number of discharges observed in 2013 FFS Medicare claims (to avoid impacts on hospital admissions and discharges from affecting the regression weights).[22,23] Giving larger hospitals more weight reflects that fact that larger hospitals will influence experience with hospital care for more beneficiaries in the state. Weighting all hospitals equally may not accurately represent the average beneficiary hospital rating in the area if ratings differ for larger versus smaller hospitals.[24]

Next, we observed a potentially problematic data anomaly in the year 2016. Several Maryland hospitals, including two of its largest (hospitals associated with the University of Maryland system and the Johns Hopkins Hospital system), did not report Hospital CAHPS scores in 2016. We do observe scores for these hospitals in all other years from 2011 to 2021. Because the missing hospitals represent a significant amount of weight in the analysis, to avoid anomalous results in 2016, we removed that year of data from our analysis for all hospitals. In addition, data from 2022 were not available for this report, though we do have Hospital CAHPS scores based on reporting from the second half of 2020, which we treat as though it were a full year of data. As such, our regression models include estimated impacts from 2014 through 2021, excluding 2016. Notably, the impacts in 2020 were consistent with impacts in earlier years and with 2021, suggesting that neither the partial year of data nor the COVID-19 pandemic appeared to have a large influence on results.

---

[22] We placed hospitals into PUMAs by geo-coding addresses using GIS software to generate X,Y coordinates for every hospital in the Hospital Compare database.

[23] If a hospital did not have claims in 2013, we used the year after the hospital appeared in our data. For example. If a hospital newly opened in 2016 we would use as their discharge weight the number of discharges in 2017.

[24] One limitation of this weighting approach is that our hospital size weights represent the average number of FFS beneficiaries (since they are based on number of discharges in the claims), but hospital ratings are for all patients.

Finally, an important key difference for estimating impacts on patients' rating of their hospital care relative to claims-based beneficiary-level and episode-level analyses is how we defined the list of covariates included in the regressions. To control for differences in case mix over time that could be the result of shifting care out of the hospital in Maryland, we included an index measuring hospital case mix based on hospital diagnosis-related groups (DRGs) in all models. We also considered other hospital-level controls, such as the hospital's wage index, measure of disproportionate share, resident-to-bed ratio, the percentage of the total population residing in a rural zip code for the PUMA the hospital is located in, and the average SVI score of the PUMA, in addition to hospital fixed effects. Our final models chose not to control for any of these additional measures (other than case mix index) or for hospital fixed effects because doing so might control for differences in hospitals that could be the result of hospitals closing and leaving our sample. To the extent that the Maryland Model supports hospitals financially in a way that would avoid closures (leading to a more stable set of hospitals over time relative to the comparison group), we aimed to capture those effects as part of our impact estimates. Ultimately, this decision did not affect our conclusions because including hospital fixed effects and the hospital controls listed above did not materially change our results.

B.1.7. Regression-adjusted means and percentage impact

To help interpret the estimated difference-in-differences impact estimates, and to help understand the magnitude of effects across outcomes that are on different scales, we calculated regression-adjusted means and percentage impact for each of our estimated outcomes.

B.1.7a. Regression-adjusted means

Regression-adjusted means help the reader decompose the difference-in-differences impact estimate into its component parts: the mean in Maryland and the mean in the comparison group, before and after the intervention. In all periods, including baseline (2011–2013), MDAPM (2014–2018), and MD TCOC (2019–2022), and their individual years, the regression-adjusted mean for Maryland is simply the mean of the outcome in Maryland during that period or year (weighted for observability in claims-based beneficiary-year analyses).

For the comparison group, in the baseline period, we calculated the regression-adjusted mean as the mean of the outcome in the comparison group weighted by the PUMA matching weights (times observability in claims-based beneficiary-year analyses). In all post-baseline years (2014–2022), we calculated the regression-adjusted mean in the comparison group as the Maryland mean in that period or year minus the difference-in-differences impact estimate associated with that period or year, minus the difference between Maryland and the comparison group in the baseline period. For example, Maryland averaged 340 all-cause admissions per 1,000 beneficiaries during our baseline, compared with 334 all-cause admissions in our weighted comparison group during that time, for a difference of 6 admissions per 1,000 beneficiaries. In 2022, Maryland's admissions had fallen to 220 per 1,000 beneficiaries. To calculate the regression-adjusted comparison group mean, we took 220, minus the estimated difference-in-differences impact of -39 admissions, minus the difference of 6 admissions from the baseline to get 253 admissions per 1,000 beneficiaries. This approach ensures that the difference between Maryland and the comparison group at baseline (first difference) minus the difference between Maryland and the comparison group in 2022 (second difference) equals the estimated impact in 2022.

B.1.7b. Percentage impact

Percentage impacts help describe the magnitude of impact estimates on a scale common to all outcomes. In all post-baseline years (2014–2022) we calculated the percentage impact as the ratio of the impact estimate in any given year to the estimated counterfactual, multiplied by 100. The estimated counterfactual is the difference between the actual Maryland mean and the estimated impact. Using the same example as above, in 2022, we calculated the percentage impact on all-cause admissions as the impact estimate of -39 divided by the difference between the Maryland mean of 220 minus the impact estimate of -39, which equals 15.1%. The percentage impact for estimates of whether the Maryland Model changed outcomes more during the MD TCOC period than it did at the end of the MDAPM period are calculated slightly differently. For these, we simply subtracted the two percentage impacts; the MD TCOC period minus the end of the MDAPM period for a percentage point difference. For example, -15.1% (MD TCOC period percentage impact) minus -10.6% (MDAPM period percentage impact) equals 5.6 percentage points.

B.1.8. Tables of impact estimates and regression adjusted means by year

In this section, we present, in tables, regression-adjusted means as well as impact estimates of the model by year (Exhibit B.13-B.15). Using all-cause admissions as an example (Exhibit B.14), the following is a description of how readers can interpret the tables in this section:

- The regression-adjusted means during the baseline period show little difference between the intervention and comparison groups in the admission rate (340 versus 334 admissions per 1,000 beneficiaries per year), as expected (and required through matching).
- From baseline (2011–2013) to the first year of MD TCOC (2019), admissions declined faster in Maryland than for the comparison group (87 [=340 – 253] versus 39 [=334 – 295] per 1,000 beneficiaries, respectively). Thus, the difference-in-differences estimate for the Maryland Model during the first year of the MD TCOC period was -48 admissions per 1,000 beneficiaries (=-87- [-39]). This is a 15.9% reduction =-48 / (253 – [-48]) with a 90% CI of -57 to -40. As reflected in the 90% CI, this estimate is statistically different from zero ($p < 0.05$).[25] We calculated the impacts in 2020, 2021, and 2022 the same way.
- Combining the four estimates from 2019, 2020, 2021, and 2022, we estimate an average effect during the four years of the MD TCOC period of -44 admissions per 1,000 beneficiaries, which is statistically significant.

[25] The percentage equals the impact estimate divided by the estimated counterfactual (which equals the Maryland mean minus the impact estimate).

- We calculated the difference in estimates during the MD TCOC period and later MDAPM period in the same way, but we used the combined later MDAPM period estimates as the baseline. Using 2022 as an example, a decline in Maryland of 50 per 1,000 beneficiaries (from 270 to 220) compared with a decline in the comparison group of 43 per 1,000 beneficiaries (from 296 to 253) represents a difference-in-differences estimate of -8 admissions per 1,000 beneficiaries, which is statistically significant ($p < 0.05$) with a 90% CI of -12 to -4.
- To calculate the change in impact from the end of the MDAPM period to the MD TCOC period, we subtracted the percent impact during the end of the MDAPM period from the estimate in the MD TCOC period. Continuing the example from above, the impact estimate for admissions is 4.5 percentage points larger in 2022 (15.1%) than the estimate during the later MDAPM period (10.6%).

See Appendix B.1.7. for regression model specification details that produce the different impact estimates and their CIs.

B.1.8a. Impacts on Medicare FFS spending

Exhibit B.13. Impacts of the Maryland Model on Medicare FFS spending, dollars per beneficiary per year

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Total Medicare FFS spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$12,268	$11,092	$1,176		
Early MDAPM period					
2014	$12,085	$11,100	$985	-$191** (-$282; -$100)	-1.6%
2015	$12,379	$11,269	$1,110	-$66 (-$157; $24)	-0.5%
2016	$12,335	$11,348	$987	-$189** (-$277; -$101)	-1.5%
Later MDAPM period					
2017	$12,745	$11,629	$1,116	-$60 (-$190; $70)	-0.5%
2018	$13,027	$12,085	$942	-$234** (-$365; -$102)	-1.8%
Combined (2017–2018)	$12,887	$11,858	$1,029	-$147* (-$272; -$23)	-1.1%
MD TCOC period					
2019	$13,401	$12,689	$712	-$464** (-$610; -$318)	-3.3%
2020	$12,885	$12,162	$723	-$453** (-$619; -$288)	-3.4%
2021	$14,148	$13,182	$966	-$210* (-$414; -$7)	-1.5%
2022	$14,639	$13,488	$1,151	-$25 (-$224; $174)	-0.2%
Combined (2019-2022)	$13,755	$12,871	$884	-$292** (-$451; -$133)	-2.1%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$316** (-$394; -$239)	-2.2pp
2020				-$306** (-$416; -$196)	-2.3pp
2021				-$63 (-$214; $89)	-0.4pp
2022				$122 (-$35; $280)	0.9pp
Combined (2019–2022)				-$145** (-$241; -$48)	-1.0pp
Total Medicare spending + non-claims payments ($ per beneficiary per year)[b]					
Baseline period (2011–2013)	$12,259	$11,088	$1,171		
Early MDAPM period					
2014	$12,084	$11,101	$983	-$188** (-$280; -$96)	-1.5%
2015	$12,395	$11,279	$1,116	-$55 (-$146; $37)	-0.4%
2016	$12,358	$11,358	$1,000	-$171** (-$258; -$84)	-1.4%
Later MDAPM period					
2017	$12,740	$11,655	$1,085	-$86 (-$215; $43)	-0.7%
2018	$13,035	$12,119	$916	-$255** (-$385; -$125)	-1.9%
Combined (2017–2018)	$12,888	$11,888	$1,000	-$171** (-$294; -$48)	-1.3%
MD TCOC period					
2019	$13,482	$12,752	$730	-$441** (-$585; -$297)	-3.2%
2020	$13,084	$12,275	$773	-$398** (-$565; -$232)	-3.0%
2021	$14,385	$13,289	$1,096	-$75 (-$276; $126)	-0.5%
Combined (2019–2021)	$13,362	$12,768	$864	-$307** (-$463; -$150)	-2.2%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$270** (-$346; -$194)	-1.9pp
2020				-$227** (-$339; -$115)	-1.7pp
2021				$96 (-$52; $244)	0.8pp
Combined (2019–2021)				-$136** (-$227; -$44)	-0.9pp
Hospital spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$6,926	$5,417	$1,509		
Early MDAPM period					
2014	$6,799	$5,490	$1,309	-$200** (-$272; -$129)	-2.9%
2015	$6,907	$5,563	$1,344	-$165** (-$235; -$95)	-2.3%
2016	$6,814	$5,659	$1,155	-$354** (-$424; -$284)	-4.9%
Later MDAPM period					
2017	$7,077	$5,830	$1,247	-$262** (-$361; -$163)	-3.6%
2018	$7,105	$6,030	$1,075	-$434** (-$531; -$336)	-5.8%
Combined (2017-2018)	$7,091	$5,930	$1,161	-$348** (-$442; -$255)	-4.7%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
MD TCOC period					
2019	$7,225	$6,305	$920	-$589** (-$702; -$475)	-7.5%
2020	$7,081	$6,030	$1,051	-$458** (-$593; -$324)	-6.1%
2021	$7,490	$6,510	$980	-$529** (-$685; -$374)	-6.6%
2022	$7,880	$6,709	$1,171	-$338** (-$485; -$191)	-4.1%
Combined (2019-2022)	$7,413	$6,384	$1,029	-$480** (-$603; -$357)	-6.1%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$240** (-$299; -$182)	-2.8pp
2020				-$110** (-$199; -$22)	-1.4pp
2021				-$181** (-$294; -$68)	-1.9pp
2022				$10 (-$102; $123)	0.6pp
Combined (2019–2022)				-$132** (-$204; -$60)	-1.4pp
Non-hospital spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$5,342	$5,675	-$333		
Early MDAPM period					
2014	$5,286	$5,610	-$324	$9 (-$28; $47)	0.2%
2015	$5,472	$5,707	-$235	$98** ($59; $138)	1.8%
2016	$5,520	$5,687	-$167	$166** ($125; $206)	3.1%
Later MDAPM period					
2017	$5,668	$5,800	-$132	$201** ($144; $259)	3.7%
2018	$5,922	$6,055	-$133	$200** ($138; $262)	3.5%
Combined (2017–2018)	$5,795	$5,927	-$132	$201** ($143; $259)	3.6%
MD TCOC period					
2019	$6,177	$6,385	-$208	$125** ($55; $195)	2.1%
2020	$5,804	$6,132	-$328	$5 (-$66; $77)	0.1%
2021	$6,658	$6,672	-$14	$319** ($235; $403)	5.0%
2022	$6,759	$6,779	-$20	$313** ($218; $408)	4.9%
Combined (2019–2022)	$6,342	$6,487	-$145	$188** ($116; $259)	3.1%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$76** (-$115; -$37)	-1.5pp
2020				-$196** (-$243; -$149)	-3.5pp
2021				$118** ($51; $185)	1.4pp
2022				$112** ($33; $191)	1.3pp
Combined (2019–2022)				-$13 (-$60; $34)	-0.5p

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Post-acute care spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$1,149	$1,189	-$40		
Early MDAPM period					
2014	$1,095	$1,151	-$56	-$16* (-$32; $0)	-1.4%
2015	$1,123	$1,162	-$39	$1 (-$18; $20)	0.1%
2016	$1,079	$1,116	-$37	$3 (-$16; $23)	0.3%
Later MDAPM period					
2017	$1,044	$1,109	-$65	-$25* (-$49; -$1)	-2.3%
2018	$1,018	$1,107	-$89	-$49** (-$77; -$22)	-4.6%
Combined (2017–2018)	$1,031	$1,108	-$77	-$37** (-$62; -$13)	-3.5%
MD TCOC period					
2019	$1,008	$1,131	-$123	-$83** (-$114; -$53)	-7.6%
2020	$962	$1,130	-$168	-$128** (-$164; -$91)	-11.7%
2021	$1,041	$1,126	-$85	-$45* (-$85; -$6)	-4.1%
2022	$1,118	$1,211	-$93	-$53** (-$97; -$9)	-4.5%
Combined (2019–2022)	$1,031	$1,149	-$118	-$78** (-$112; -$44)	-7.0%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$46** (-$61; -$30)	-4.1pp
2020				-$90** (-$117; -$64)	-8.2pp
2021				-$8 (-$38; $22)	-0.6pp
2022				-$15 (-$54; $23)	-1.0pp
Combined (2019–2022)				-$40** (-$64; -$17)	-3.5pp
Ambulatory care visit with primary care providers and specialist physicians spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$632	$604	$28		
Early MDAPM period					
2014	$656	$624	$32	$4** ($1; $7)	0.6%
2015	$685	$640	$45	$17** ($13; $21)	2.5%
2016	$694	$646	$48	$20** ($16; $24)	3.0%
Later MDAPM period					
2017	$715	$661	$54	$26** ($22; $31)	3.8%
2018	$742	$683	$59	$31** ($26; $37)	4.4%
Combined (2017–2018)	$728	$671	$57	$29** ($24; $34)	4.1%
MD TCOC period					
2019	$784	$722	$62	$34** ($27; $40)	4.5%
2020	$691	$649	$42	$14** ($7; $22)	2.1%
2021	$936	$844	$92	$64** ($50; $79)	7.3%
2022	$948	$861	$87	$59** ($44; $74)	6.6%
Combined (2019–2022)	$838	$768	$70	$42** ($32; $52)	5.3%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Difference in estimates during MD TCOC period and later MDAPM period					
2019				$5** ($1; $9)	0.4pp
2020				-$14** (-$20; -$9)	-2.0pp
2021				$35** ($23; $48)	3.2pp
2022				$30** ($17; $43)	2.5pp
Combined (2019–2022)				$14** ($6; $21)	1.2pp
Non-hospital Part B drug spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$318	$341	-$23		
Early MDAPM period					
2014	$354	$363	-$9	$14** ($2; $25)	4.1%
2015	$380	$382	-$2	$21** ($7; $35)	5.8%
2016	$427	$418	$9	$32** ($11; $52)	8.1%
Later MDAPM period					
2017	$486	$450	$36	$59** ($36; $83)	13.8%
2018	$568	$511	$57	$80** ($51; $109)	16.4%
Combined (2017–2018)	$527	$480	$47	$70** ($44; $95)	15.3%
MD TCOC period					
2019	$643	$584	$59	$82** ($49; $116)	14.6%
2020	$684	$611	$73	$96** ($59; $133)	16.3%
2021	$800	$688	$112	$135** ($91; $179)	20.3%
2022	$847	$734	$113	$136** ($89; $183)	19.1%
Combined (2019–2022)	$742	$653	$89	$112** ($73; $150)	17.8%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				$13 (-$4; $29)	-0.7pp
2020				$27** ($7; $46)	1.0pp
2021				$65** ($35; $95)	5.0pp
2022				$66** ($31; $101)	3.8pp
Combined (2019–2022)				$42** ($20; $65)	2.5pp
Ambulatory surgical center facility spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$127	$87	$40		
Early MDAPM period					
2014	$133	$92	$41	$1 (-$1; $4)	0.8%
2015	$137	$96	$41	$1 (-$2; $4)	0.7%
2016	$145	$103	$42	$2 (-$1; $6)	1.4%
Later MDAPM period					
2017	$153	$109	$44	$4** ($1; $7)	2.7%
2018	$168	$117	$51	$11** ($6; $15)	7.0%
Combined (2017–2018)	$161	$114	$47	$7** ($4; $11)	4.5%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
MD TCOC period					
2019	$185	$127	$58	$18** ($13; $23)	10.8%
2020	$170	$118	$52	$12** ($7; $17)	7.6%
2021	$216	$146	$70	$30** ($24; $37)	16.1%
2022	$233	$160	$73	$33** ($26; $40)	16.5%
Combined (2019–2022)	$200	$137	$63	$23** ($18; $28)	13.0%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				$11** ($7; $14)	6.3pp
2020				$5* ($1; $9)	3.1pp
2021				$23** ($17; $29)	11.6pp
2022				$26** ($19; $33)	12.0pp
Combined (2019–2022)				$16** ($12; $21)	8.5pp
Imaging and testing professional spending ($ per beneficiary per year)					
Baseline period (2011–2013)	$695	$612	$83		
Early MDAPM period					
2014	$648	$582	$66	-$17** (-$24; -$10)	-2.6%
2015	$656	$575	$81	-$2 (-$10; $7)	-0.3%
2016	$662	$568	$94	$11** ($2; $19)	1.7%
Later MDAPM period					
2017	$688	$586	$102	$19** ($10; $28)	2.8%
2018	$721	$610	$111	$28** ($19; $37)	4.0%
Combined (2017–2018)	$705	$598	$107	24** ($15; $33)	3.5%
MD TCOC period					
2019	$753	$633	$120	$37** ($27; $48)	5.2%
2020	$688	$600	$88	$5 (-$9; $20)	0.7%
2021	$845	$690	$155	$72** ($60; $84)	9.3%
2022	$814	$691	$123	$40** ($26; $54)	5.2%
Combined (2019–2022)	$200	$137	$63	$39** ($27; $50)	5.3%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				$13** ($8; $19)	1.7pp
2020				-$18** (-$29; -$8)	-2.8pp
2021				$49** ($38; $59)	5.8pp
2022				$16** ($3; $29)	1.7pp
Combined (2019–2022)				$15** ($7; $23)	1.8pp

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Other non-hospital ($ per beneficiary per year)					
Baseline period (2011–2013)	$2,422	$2,842	-$420		
Early MDAPM period					
2014	$2,401	$2,797	-$396	$24* ($2; $45)	1.0%
2015	$2,492	$2,852	-$360	$60** ($35; $85)	2.5%
2016	$2,513	$2,835	-$322	$98** ($69; $126)	4.1%
Later MDAPM period					
2017	$2,581	$2,883	-$302	$118** ($80; $156)	4.8%
2018	$2,705	$3,026	-$321	$99** ($59; $140)	3.8%
Combined (2017–2018)	$2,644	$2,955	-$311	$109** ($70; $147)	4.3%
MD TCOC period					
2019	$2,804	$3,187	-$383	$37 (-$7; $81)	1.3%
2020	$2,610	$3,026	-$416	$4 (-$41; $50)	0.2%
2021	$2,820	$3,178	-$358	$62** ($12; $112)	2.2%
2022	$2,799	$3,121	-$322	$98** ($47; $149)	3.6%
Combined (2019–2022)	$2,757	$3,127	-$370	$50* ($6; $93)	1.8%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$72** (-$93; -$51)	-3.0pp
2020				-$104** (-$128; -$80)	-4.1pp
2021				-$46** (-$78; -$14)	-2.1pp
2022				-$11 (-$47; $25)	-0.7pp
Combined (2019–2022)				-$59** (-$81; -$37)	-2.5pp

* $p < 0.10$; ** $p < 0.05$.

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

[b] Total Medicare spending + non-claims payments in 2021 was not available for this report.

CI = confidence interval ; pp = percentage points.

B.1.8b. Impacts on health care utilization and quality

Exhibit B.14. Impacts of the Maryland Model on health care utilization

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
All-cause admissions (number per 1,000 beneficiaries per year)					
Baseline period (2011–2013)	340	334	6		
Early MDAPM period					
2014	303	305	-2	-8** (-13; -3)	-2.6%
2015	292	302	-10	-16** (-22; -10)	-5.2%
2016	283	296	-13	-19** (-25; -13)	-6.3%
Later MDAPM period					
2017	275	297	-22	-28** (-35; -21)	-9.2%
2018	265	294	-29	-35** (-43; -28)	-11.7%
Combined (2017-2018)	270	296	-26	-32** (-38; -25)	-10.6%
MD TCOC period					
2019	253	295	-42	-48** (-57; -40)	-15.9%
2020	214	256	-42	-48** (-56; -40)	-18.3%
2021	221	254	-33	-39** (-47; -31)	-15.0%
2022	220	253	-33	-39** (-48; -31)	-15.1%
Combined (2019-2022)	227	265	-38	-44** (-52; -36)	-16.2%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-17** (-20; -13)	-5.3pp
2020				-16** (-21; -12)	-7.7pp
2021				-7** (-12; -3)	-4.4pp
2022				-8** (-12; -4)	-4.5pp
Combined (2019-2022)				-12** (-16; -8)	-5.6pp
Outpatient emergency department and observation stays (number per 1,000 beneficiaries per year)					
Baseline period (2011-2013)	456	448	8		
Early MDAPM period					
2014	478	473	5	-3 (-9; 4)	-0.6%
2015	487	486	1	-7 (-15; 1)	-1.4%
2016	475	489	-14	-22** (-29; -14)	-4.4%
Later MDAPM period					
2017	482	490	-8	-16** (-24; -8)	-3.2%
2018	480	485	-5	-13** (-22; -4)	-2.6%
Combined (2017–2018)	481	487	-6	-14** (-22; -6)	-2.8%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
MD TCOC period					
2019	475	484	-9	-17** (-27; -8)	-3.5%
2020	355	371	-16	-24** (-32; -16)	-6.3%
2021	385	396	-11	-19** (-29; -10)	-4.7%
2022	384	414	-30	-38** (-47; -28)	-9.0%
Combined (2019–2022)	400	417	-17	-25** (-33; -16)	-5.9%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-3 (-8; 1)	-0.7pp
2020				-10** (-17; -3)	-3.5pp
2021				-5 (-13; 2)	-1.9pp
2022				-23** (-32; -15)	-6.2pp
Combined (2019–2022)				-10** (-16; -4)	-3.1pp
Intensity of hospital care (measured by standardized hospital spending) ($ per beneficiary per year)[b]					
Baseline period (2011–2013)	$4,564	$4,585	-$21		
Early MDAPM period					
2014	$4,839	$4,651	$188	$209** ($136; $282)	4.5%
2015	$4,727	$4,746	-$19	$2 (-$47; $51)	0.0%
2016	$4,665	$4,813	-$148	-$127** (-$177; -$77)	-2.7%
Later MDAPM period					
2017	$4,804	$4,961	-$157	-$136** (-$196; -$76)	-2.8%
2018	$4,817	$5,163	-$346	-$325** (-$390; -$260)	-6.3%
Combined (2017–2018)	$4,810	$5,062	-$252	-$231** (-$290; -$172)	-4.6%
MD TCOC period					
2019	$4,882	$5,341	-$459	-$438** (-$513; -$364)	-8.2%
2020	$4,583	$5,016	-$433	-$412** (-$498; -$327)	-8.2%
2021	$4,958	$5,389	-$431	-$410** (-$513; -$307)	-7.6%
Combined (2019–2021)	$4,806	$5,247	-$441	-$420** (-$502; -$339)	-8.0%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-$208** (-$250; -$165)	-3.6pp
2020				-$181** (-$241; -$122)	-3.6pp
2021				-$179** (-$257; -$101)	-3.0pp
Combined (2019–2021)				-$189** (-$241; -$138)	-3.4pp
Potentially preventable admissions (number per 1,000 beneficiaries per year)					
Baseline period (2011-2013)	60.6	60.4	0.2		
Early MDAPM period					
2014	53.6	53.5	0.1	-0.1 (-1.3; 1.1)	-0.2%
2015	52.8	52.7	0.1	-0.1 (-1.6; 1.3)	-0.2%
2016	49.4	50.2	-0.8	-1.0 (-2.5; 0.5)	-2.0%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Later MDAPM period					
2017	49.2	52.8	-3.6	-3.8** (-5.6; -2.0)	-7.2%
2018	45.5	51.9	-6.4	-6.6** (-9.1; -4.0)	-12.7%
Combined (2017-2018)	47.3	52.3	-5.0	-5.2** (-7.3; -3.1)	-9.9%
MD TCOC period					
2019	44.2	51.8	-7.6	-7.8** (-10.3; -5.4)	-15.0%
2020	33.2	40.1	-6.9	-7.1** (-9.4; -4.7)	-17.6%
2021	32.3	39.3	-7.0	-7.2** (-9.8; -4.6)	-18.2%
2022	32.6	39.1	-6.5	-6.7** (-9.3; -4.1)	-17.0%
Combined (2019-2022)	35.6	42.6	-7.0	-7.2** (-9.6; -4.8)	-16.8%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-2.6** (-3.5; -1.7)	-5.1pp
2020				-1.9** (-3.1; -0.6)	-7.7pp
2021				-2.0** (-3.3; -0.8)	-8.3pp
2022				-1.5* (-2.8; -0.1)	-7.1pp
Combined (2019-2022)				-2.0** (-3.0; -1.0)	-6.9pp
30-day post-discharge unplanned readmissions (percentage of discharges)					
Baseline period (2011-2013)	19.1%	18.4%	0.7pp		
Early MDAPM period					
2014	17.9%	17.4%	0.5pp	-0.2pp (-0.5pp; 0.0pp)	-1.1%
2015	17.1%	17.0%	0.1pp	-0.6pp** (-0.9pp; -0.4pp)	-3.4%
2016	16.6%	16.9%	-0.3pp	-1.0pp** (-1.3pp; -0.8pp)	-5.7%
Later MDAPM period					
2017	16.3%	17.1%	-0.8pp	-1.5pp** (-1.8pp; -1.1pp)	-8.4%
2018	16.4%	17.0%	-0.6pp	-1.3pp** (-1.6pp; -1.0pp)	-7.3%
Combined (2017-2018)	16.3%	17.0%	-0.7pp	-1.4pp** (-1.7pp; -1.1pp)	-7.9%
MD TCOC period					
2019	16.0%	17.2%	-1.2pp	-1.9pp** (-2.2pp; -1.6pp)	-10.6%
2020	16.1%	17.2%	-1.1pp	-1.8pp** (-2.2pp; -1.5pp)	-10.1%
2021	16.6%	17.2%	-0.6pp	-1.3pp** (-1.6pp; -1.0pp)	-7.3%
2022	16.6%	17.2%	-0.6pp	-1.3pp** (-1.7pp; -1.0pp)	-7.3%
Combined (2019-2022)	16.3%	17.2%	-0.9pp	-1.6pp** (-1.9pp; -1.3pp)	-8.9%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-0.5pp** (-0.8pp; -0.3pp)	-2.7pp
2020				-0.4pp** (-0.7pp; -0.1pp)	-2.2pp
2021				0.0pp (-0.2pp; 0.3pp)	0.6pp
2022				0.0pp (-0.3pp; 0.4pp)	0.6pp
Combined (2019-2022)				-0.2pp* (-0.4pp; 0.0pp)	-1.0pp

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Timely follow-up after acute exacerbation of chronic conditions (percentage of discharges)					
Baseline period (2011-2013)	65.5%	65.4%	0.1pp		
Early MDAPM period					
2014	66.5%	66.7%	-0.2pp	-0.3pp (-0.7pp; 0.2pp)	-0.4%
2015	68.5%	67.2%	1.3pp	1.2pp** (0.7pp; 1.7pp)	1.8%
2016	70.6%	68.8%	1.8pp	1.7pp** (1.0pp; 2.3pp)	2.5%
Later MDAPM period					
2017	71.1%	69.4%	1.7pp	1.6pp** (0.9pp; 2.2pp)	2.3%
2018	71.3%	69.8%	1.5pp	1.4pp** (0.7pp; 2.1pp)	2.0%
Combined (2017-2018)	71.2%	69.6%	1.6pp	1.5pp** (0.9pp; 2.1pp)	2.2%
MD TCOC period					
2019	71.8%	69.9%	1.9pp	1.8pp** (1.1pp; 2.4pp)	2.6%
2020	68.6%	66.7%	1.9pp	1.8pp** (1.1pp; 2.6pp)	2.7%
2021	70.5%	68.7%	1.8pp	1.7pp** (1.0pp; 2.4pp)	2.5%
2022	71.2%	69.0%	2.2pp	2.1pp** (1.2pp; 2.9pp)	3.0%
Combined (2019-2022)	70.6%	68.7%	1.9pp	1.8pp** (1.2pp; 2.5pp)	2.6%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				0.3pp (-0.2pp; 0.8pp)	0.4pp
2020				0.3pp (-0.3pp; 1.0pp)	0.5pp
2021				0.2pp (-0.3pp; 0.7pp)	0.3pp
2022				0.6pp (-0.1pp; 1.2pp)	0.8pp
Combined (2019-2022)				0.3pp (-0.1pp; 0.8pp)	0.4pp

* $p < 0.10$; ** $p < 0.05$

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

[b] Standardized spending is included under “utilization” because it is an aggregate measure of intensity of hospital services, inpatient and outpatient. It removes differences in hospital spending between Maryland and the comparison group because of Health Services Cost Review Commission rate setting and other adjustments. Standardized hospital spending in 2021 was not available for this report.

CI = confidence interval; pp = percentage point.

B.1.8c. Impacts on patient experience and use of Medicare Diabetes Prevention Program services

Exhibit B.15. Impacts of the Maryland Model on patient experience and population health

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Patients' rating of their personal doctor (mean rating on a scale of 0 to 100)[b]					
Baseline period (2011-2013)	90.2	90.3	-0.1		
Early MDAPM period					
2014	90.5	90.0	0.5	0.6 (-0.2; 1.3)	0.7%
2015	88.9	90.0	-1.1	-1.0 (-2.1; 0.0)	-1.1%
2016	89.6	90.4	-0.8	-0.7 (-1.5; 0.2)	-0.8%
Later MDAPM period					
2017	90.4	90.3	0.1	0.2 (-0.7; 1.1)	0.2%
2018	90.3	90.4	-0.1	0.0 (-0.8; 0.8)	-0.0%
Combined (2017-2018)	90.4	90.4	0.0	0.1 (-0.5; 0.7)	0.1%
MD TCOC period					
2019	90.9	90.7	0.2	0.3 (-0.3; 1.0)	0.3%
2020	n.a.	n.a.	n.a.	n.a.	n.a.
2021	91.3	90.9	0.4	0.5 (-0.3; 1.4)	0.6%
2022	90.4	90.5	-0.1	0.0 (-1.0; 0.9)	-0.0%
Combined (2019-2022)	90.9	90.7	0.2	0.3 (-0.3; 0.9)	0.3%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				0.3 (-0.6; 1.1)	0.2pp
2020					n.a.
2021				0.5 (-0.3; 1.2)	0.5pp
2022				-0.1 (-1.1; 0.8)	-0.1pp
Combined (2019-2022)				0.2 (-0.4; 0.8)	0.2pp
Patients' rating of their hospital (percent of respondents within each hospital who rated the hospital 9 or 10 out of 10)[c]					
Baseline period (2011-2013)	66.6%	67.8%	-1.2pp		
Early MDAPM period					
2014	67.0%	68.0%	-1.0pp	0.2pp (-0.9pp; 1.4pp)	0.3%
2015	67.9%	68.3%	-0.4pp	0.8pp (-0.6pp; 2.2pp)	1.2%
2016[d]	n.a.	n.a.	n.a.	n.a.	n.a.
Later MDAPM period					
2017	68.9%	69.0%	-0.1pp	1.1pp (-0.3pp; 2.4pp)	1.6%
2018	68.3%	68.6%	-0.3pp	0.9pp (-0.6pp; 2.4pp)	1.3%
Combined (2017-2018)	68.6%	68.8%	-0.2pp	1.0pp (-0.4pp; 2.3pp)	1.5%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
MD TCOC period					
2019	68.0%	69.0%	-1.0pp	0.2pp (-1.2pp; 1.6pp)	0.3%
2020	67.8%	68.3%	-0.5pp	0.7pp (-1.0pp; 2.5pp)	1.0%
2021	66.0%	66.6%	-0.6pp	0.6pp (-1.2pp; 2.4pp)	0.9%
Combined (2019-2021)	67.3%	68.0%	-0.7pp	0.5pp (-1.0pp; 2.0pp)	0.7%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-0.8pp (-1.5pp; 0.0pp)	-1.2pp
2020				-0.2pp (-1.4pp; 0.9pp)	-0.5pp
2021				-0.4pp (-1.6pp; 0.8pp)	-0.6pp
Combined (2019-2021)				-0.5pp (-1.3pp; 0.4pp)	-0.8pp
Use of Medicare Diabetes Prevention Program services (number using services in a year per 10,000 beneficiaries)[d]					
Baseline period (2011-2013)	0.0	0.0	0.0		
Early MDAPM period					
2014	0.0	0.0	0.0	0.0 (0.0; 0.0)	n.a.
2015	0.0	0.0	0.0	0.0 (0.0; 0.0)	n.a.
2016	0.0	0.0	0.0	0.0 (0.0; 0.1)	n.a.
Later MDAPM period					
2017	0.0	0.1	-0.1	-0.1 (-0.1; 0.0)	-100.0%[e]
2018	1.3	3.1	-1.8	-1.8** (-2.7; -0.9)	-58.1%[e]
Combined (2017-2018)	0.7	1.6	-0.9	-0.9** (-1.4; -0.5)	-56.2%[e]
MD TCOC period					
2019	0.3	1.1	-0.8	-0.8* (-1.6; 0.0)	-72.7%[e]
2020	0.3	0.4	-0.1	-0.1 (-0.3; 0.2)	-25.0%[e]
2021	0.3	0.5	-0.2	-0.2 (-0.4; 0.0)	-40.0%[e]
2022	0.8	0.5	0.3	0.3* (0.0; 0.6)	60.0%[e]
Combined (2019-2022)	0.4	0.6	-0.2	-0.2 (-0.5; 0.1)	-33.3%[e]
Difference in estimates during MD TCOC period and later MDAPM period					
2019				0.1 (-0.6; 0.9)	-16.5pp
2020				0.8** (0.3; 1.4)	31.2pp
2021				0.8** (0.3; 1.2)	16.2pp
2022				1.2** (0.7; 1.8)	116.2pp
Combined (2019-2021)				0.7* (0.3; 1.2)	22.9pp

* $p < 0.10$; ** $p < 0.05$.

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

[b] Complete information on patients' rating of their personal doctor in 2020 and 2021 were not available for this report.

[c] Data on hospital ratings in 2016 were missing for several important Maryland hospitals. The analysis omits 2016 data for all hospitals to avoid spurious findings in that year.

[d] Impact estimates from 2014-2016 on use of Diabetes Prevention Program services are effectively zero because use of these services from 2014-2016 was nearly zero, before the program was rolled out and reimbursed nationally. We do not report impacts during these years. For the same reason, baseline estimates of the mean are effectively zero, but we show means in all years in the table for completeness.

[e] Percentage impacts for use of the Diabetes Prevention Program services are large even though impacts are small because the denominator used in calculating the percentage impacts (Maryland mean) was very small as well, making this number unstable, and possibly misleading. For this reason, we suppressed the calculation of the percentage impact in the main tables of this report.

CI = confidence interval; n.a. = not applicable.

B.1.9. Sensitivity analyses

B.1.9a. Controlling for health conditions measured in Medicare claims

Our main impact estimates include controls for time-varying health conditions in the regressions to limit the potential for the higher rates of Medicare Advantage enrollment in the comparison group versus Maryland to bias estimates of model impacts. In this section, we detail our rationale for including these controls, and report sensitivity results for key outcomes that remove controls for health conditions. Further empirical support for the notion that failing to control for health condition controls could cause bias in our estimates is provided in the Quantitative-Only Report for the Model's First Three Years, Appendix A.6.1. (Rotter et al. 2022).

Rationale for including health condition controls

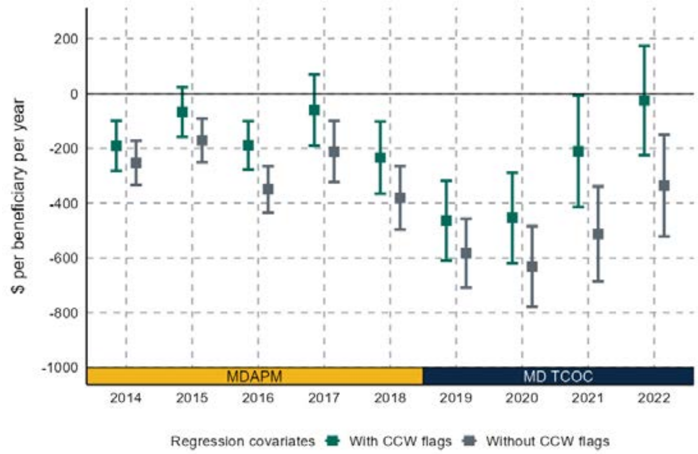
Estimating impacts of a policy or intervention requires researchers to consider all the ways the intervention might affect outcomes. Typically, when evaluating payment reform models such as the Maryland Model, we might consider changes in health status (as measured by individual health conditions) to be one of the mechanisms through which the model could improve outcomes. For example, if the model prompted primary care providers to better identify and treat early heart disease, the model could prevent some hospitalizations because of more serious heart conditions. In that case, we would not want to control for the time-varying prevalence of serious heart conditions after the intervention began because those controls could inadvertently remove some of the effects of the intervention itself. In Maryland, however, Medicare Advantage enrollment is lower, leading to a more consistent FFS Medicare population over time relative to a nationally drawn comparison group. These differences in the analytic sample we use to estimate impacts could lead to bias if, for example, healthier beneficiaries were more likely to leave for Medicare Advantage. This greater exit of healthier beneficiaries to Medicare Advantage in the comparison group would make the remaining beneficiaries in the comparison group look sicker relative to those in Maryland. This difference could, artificially, make it look like the model is making beneficiaries healthier, when really the difference in health status is just because of who is exiting to Medicare Advantage. In that case, we *would* want to control for time-varying prevalence of conditions to account for changes in the population that are unrelated to the effects of the model itself.

To estimate impacts of the model, we are balancing the concern of over-controlling for model impacts with the threat of bias that stems from higher rates of Medicare Advantage enrollment in the intervention group than in the comparison group.

Results of sensitivity analysis that removes health condition controls

- In general, we find that impacts on key outcomes, particularly spending outcomes, were qualitatively consistent but moderately larger when we removed health condition controls (Exhibits B.16 and B.17).
- The difference between models that did and did not control for health conditions was largest in 2021 and 2022. In general, the difference is growing over time, which is consistent with the idea that the population in Maryland and the comparison group continue to diverge on health status because of differential enrollment in Medicare Advantage.
- Impacts on utilization and quality of care outcomes such as all-cause, acute-care hospitalizations and preventable hospitalizations were only minorly affected by the decision to include or not include chronic health conditions (Exhibit B.17).

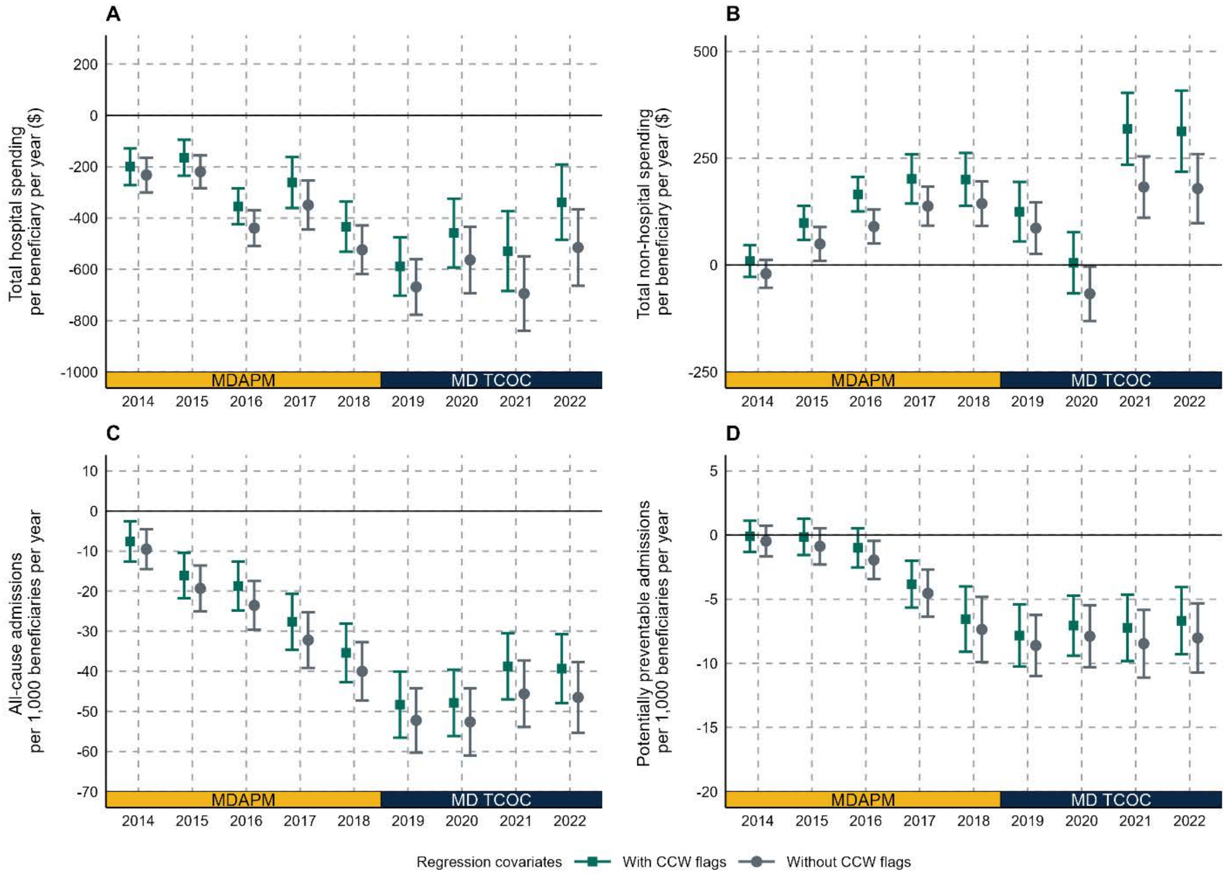
Exhibit B.16 Estimated impact of the model on total Medicare spending with and without time-varying health condition controls, by year



Note: Errors bars are 90% CIs for the yearly impact estimates. Estimates in which the intervals do not span zero are statistically different from zero at a $p < 0.10$ threshold.

CCW = Chronic Conditions Data Warehouse; CI = confidence interval.

Exhibit B.17. Impact of the Maryland Model on key measures with and without time-varying health condition controls, by year



Note: Errors bars are 90% CIs for the yearly impact estimates. Estimates in which the intervals do not span zero are statistically different from zero at a $p < 0.10$ threshold.

CI = confidence interval.

B.1.9b Accounting for the COVID-19 pandemic in the impact estimates

In this section, we describe how the COVID-19 pandemic could bias our impact estimates from 2020-2022 and the actions we took to help mitigate identified risks.

The COVID-19 pandemic could introduce bias in our results if the pandemic affected outcomes in Maryland and our selected comparison group differently in ways not related to the Maryland Model. The bias could also occur if Medicare beneficiaries, including those who do not get COVID-19, respond differently (for example, if beneficiaries in Maryland are more or less likely to avoid hospital care because of the COVID-19 pandemic). On the other hand, it's possible that the model had a true effect on our key outcomes by affecting COVID-19 related outcomes. For example, MDPCP might have helped practices learn about COVID-19 early, or the model might have allowed hospitals more flexibility and financial security under global budgets that improved access for Maryland beneficiaries (Perman et al. 2021). Indeed, the percent of the population with at least one dose of the COVID-19 vaccine as of 12/31/2021 was higher in Maryland

(80.6%) than it was nationally (73.5%) according to the Centers for Disease Control and Prevention (CDC) (2022). We do not want to adjust away any true effects the model might have on the rate of COVID-19 or COVID-19-related outcomes in Maryland. For this reason, our primary regression models *do not* include controls for COVID-19 hospitalizations and emergency department (ED) visits.

To mitigate the risk of bias in our estimates from the COVID-19 pandemic, we took a multipronged approach, including accounting for social vulnerability in matching and through regression specifications and sensitivity analyses.

Accounting for social vulnerability in matching and checking balance on COVID-19 outcomes

We chose not to include COVID-19 variables in our matching to avoid matching on future outcomes the model might have the ability to affect. We did, however, include in our matching several of the individual components from the CDC's SVI, with the idea that we want Maryland and our comparison groups to have similar levels of vulnerability to disease outbreaks, including the COVID-19 pandemic. Specifically, we matched on the following variables (defined at the PUMA level): the percentage of the population living in multi-unit structures, mobile homes, or group quarters; the percentage older than 64; the percentage younger than 18; the percentage with a high school degree (or equivalent); the percentage that speaks English well; the percentage living in a crowded home; and the percentage without a vehicle. Together with other matching variables, we captured most components of the SVI that enabled us to find a comparison group with a similar level of social vulnerability as Maryland (in 2011–2013).

We also checked balance (without including it directly in our matching algorithm) on 2020 and 2021 (the height of the COVID-19 pandemic) COVID-19 measures in Maryland and our selected comparison group (Exhibit B.18). We found the following:

- Rates of ED and observation visits for COVID-19 in Maryland and the selected comparison group were broadly similar in 2020 and very similar in 2021, with a weighted difference of 0.14 fewer visits per 1,000 people in Maryland than in the comparison group in 2021.
- Rates of COVID-19 hospitalizations were similar between groups in 2021, with a weighted difference of 1.14 fewer stays per 1,000 people in Maryland than in the comparison group. In terms of standardized differences, the number of COVID-19 hospitalizations in 2020 was lower in Maryland than in our selected comparison group. But the size of this difference was small relative to all hospital admissions (a difference of about five hospitalizations per 1,000 people was about 1.6% of total inpatient hospitalizations in Maryland in 2013).
- The declines from 2019 to 2020 in hospitalizations (all-cause, elective, and surgical) and outpatient ED visits were similar between Maryland and the selected comparison group. This indicates that the large declines in service use that occurred early in the COVID-19 pandemic occurred in similar amounts in Maryland and the comparison group.

Exhibit B.18. Balance on COVID-19 and COVID-19-related variables

Variable description	Maryland mean	Comparison group pre-weighted mean	Difference pre-weighting	Standardized difference pre-weighting	Included in matching?	Comparison group post-weighted mean	Difference post-weighting	Standardized difference post-weighting
Percentage living in multi-unit structure, mobile home, or group quarters	19.04	25.60	-6.56	-0.48	Yes	23.55	-4.52	-0.33
Percentage older than age 64	13.86	14.97	-1.11	-0.26	Yes	14.34	-0.48	-0.11
Percentage younger than age 18	22.52	22.84	-0.32	-0.09	Yes	23.21	-0.69	-0.20
Percentage with high school degree (or equivalent)	92.06	89.65	2.41	0.44	Yes	91.16	0.90	0.17
Percentage that speaks English well	97.57	96.43	1.14	0.25	Yes	96.88	0.69	0.15
Percentage living in crowded home	6.45	9.00	-2.56	-0.47	Yes	8.04	-1.59	-0.29
Percentage without a vehicle	7.03	6.34	0.69	0.09	Yes	7.36	-0.33	-0.04
Number of COVID-19 hospitalizations (2020)	13.66	16.50	-2.84	-0.34	No	18.68	-5.02	-0.59
Number of COVID-19 hospitalizations (2021)	13.85	15.63	-1.78	NA	No	13.43	-1.14	NA
Number of COVID-19 outpatient ED visits and observation stays (2020)	6.40	9.07	-2.67	-0.51	No	7.56	-1.16	-0.22
Number of COVID-19 outpatient ED visits and observation stays (2021)	10.30	12.30	-2.00	NA	No	10.56	0.14	NA
Excess number of all-cause acute care hospital admissions (2020 minus 2019)[a]	-39.69	-39.95	0.26	0.02	No	-40.49	0.80	0.06
Excess number of outpatient ED visits and observation stays (2020 minus 2019)[a]	-121.10	-114.80	-6.31	-0.19	No	-117.19	-3.91	-0.12
Excess number of ED visits and observation stays ending in inpatient stay (2020 minus 2019)[a]	-27.08	-24.16	-2.92	-0.23	No	-25.36	-1.72	-0.14
Number of surgical hospitalizations (2020 minus 2019)[a]	-14.75	-14.10	-0.65	-0.14	No	-14.22	-0.52	-0.12
Number of elective hospitalizations (2020 minus 2019)[a]	-11.46	-11.73	0.27	0.05	No	-11.31	-0.15	-0.03

[a] Difference in the 2020 rate per 1,000 Medicare fee-for-service beneficiaries and the 2019 rate per 1,000 Medicare fee-for-service beneficiaries.

ED = emergency department; NA = not available.

Regression-based approaches to account for the COVID-19 pandemic

In addition to matching, we implemented a few regression-based mitigation strategies related to the COVID-19 pandemic in our main regression specification and through sensitivity analyses.

- First, we designed our regression models to estimate the combined effect of the model from 2019 through 2022, as well as the individual yearly effects separately. Doing so allows us to interpret the effect of the model separately in its first four years. If we see large differences between yearly estimates that we think are unlikely to be related to changes made to the model, we likely would interpret those differences as attributable, at least in part, to the direct or indirect effects of the COVID-19 pandemic. As shown in Exhibits B.19 and B.20, the impact estimates were similar early in the pandemic (2020 and 2021) to what they were in 2019 for most outcomes.
- Second, each of our regression models explicitly control for the SVI measure noted above. The measure itself represents a percentile ranking of vulnerability (which is different from the individual components we included in matching) and is designed to further control for differences between Maryland and the comparison group on social vulnerability.
- Third, we conducted a sensitivity analysis in which we included a flag for COVID-19 hospitalizations and ED visits in our regression models. If we believe that COVID-19 is largely exogenous (that is, not influenced by the model), these estimates will control for differences between Maryland and the comparison group that we should otherwise not be attributing to the model.

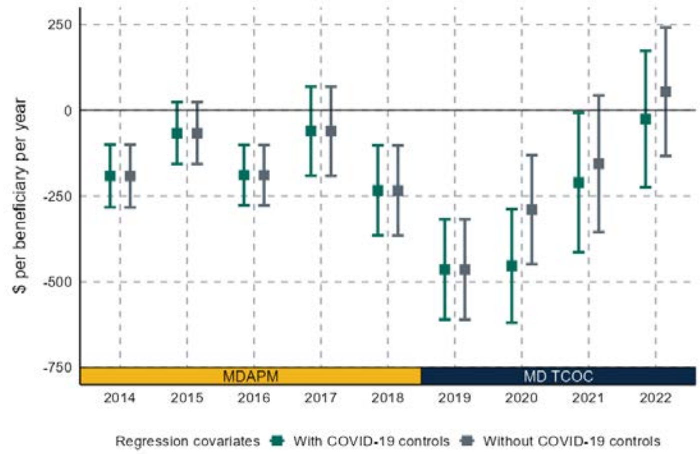
We ran sensitivity analyses for our each of our key outcomes that include as covariates COVID-19 hospitalizations and ED visits in from 2020-2022 to control for differences in the rate of these outcomes between Maryland and our comparison group. Specifically, we add as a control variable for each yearly observation whether a beneficiary had a hospital visit (inpatient or outpatient ED) with a COVID-19 diagnosis that year. In general, controlling for COVID-19 outcomes led to impact estimates that were closer to zero from 2020-2022, especially for spending outcomes. In all cases, though the impact estimates were smaller, qualitative conclusions (including statistically significant findings) did not change, and the estimates became more similar the further from the initial COVID-19 pandemic year (e.g., 2022).

Several recent articles have argued (Haft et al. 2020; Peterson and Schumacher 2020) that the Maryland Model—including MDPCP and hospital global budgets—might have decreased the rates and severity of COVID-19 in the state and improved care for patients with COVID-19. These articles suggest that controlling for COVID-19 rates is inappropriate because it could control away effects of the program. Because of this, we believe the main regression specification—which does not control for COVID-19 hospital visits—is the most appropriate. Nonetheless, our main results and conclusions are not sensitive to adding the COVID-19 controls.

Results of sensitivity analysis that controls for COVID-19 hospitalizations and ED visits

- Impacts on key outcomes with and without COVID-19 controls were very similar from 2020-2022, particularly for service use and quality outcomes (Exhibits B.19 and B.20).
- The largest differences came during the first year of the COVID-19 pandemic (2020) for hospital and total Medicare spending. Controlling for COVID-19 controls led to impacts on hospital and total spending that were modestly larger in 2020, but the substantive conclusions were the same.

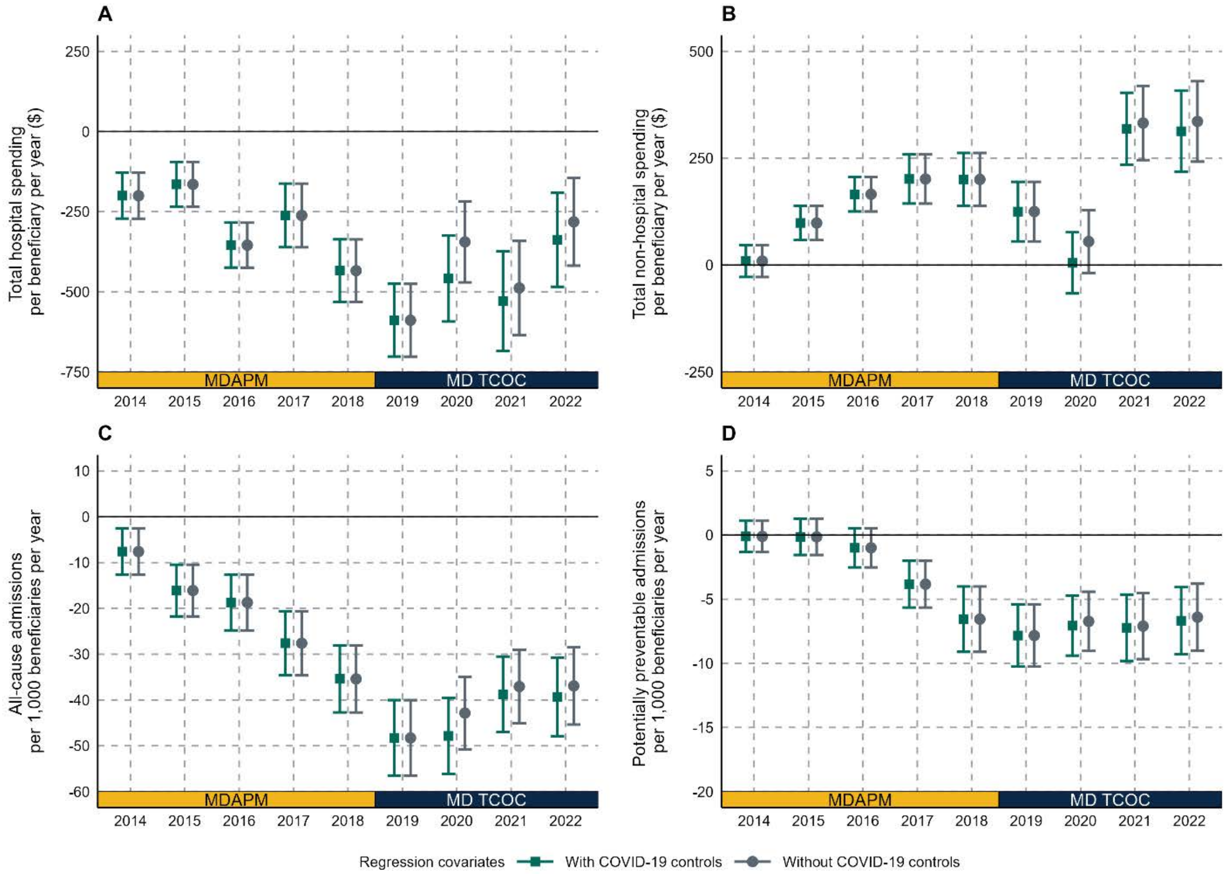
Exhibit B.19. Estimated impact of the model on total Medicare spending with and without COVID controls, by year



Note: Errors bars are 90% CIs for the yearly impact estimates. Estimates in which the intervals do not span zero are statistically different from zero at a p < 0.10 threshold.

CCW = Chronic Conditions Data Warehouse; CI = confidence interval.

Exhibit B.20. Impact of the model on key measures with and without COVID controls, by year



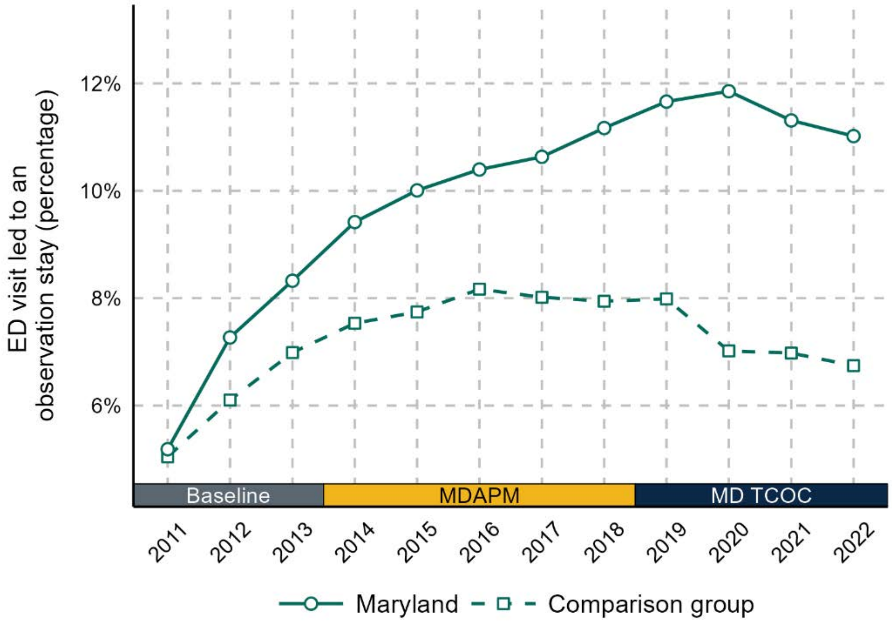
Note: Errors bars are 90% CIs for the yearly impact estimates. Estimates in which the intervals do not span zero are statistically different from zero at a $p < 0.10$ threshold.

CI = confidence interval.

B.1.10. ED discharge destination: hospital admission, hospital observation use, and discharge to community

Hospitals use observation stays as an alternative to inpatient admissions when the patient requires extra monitoring, but the hospital care team has not yet decided to admit them. In Maryland, observation stay use increased faster than it did the nation from 2011-2019 (Exhibit B.21), including during baseline (2011-2013).

Exhibit B.21. Unadjusted likelihood of being sent to observation stay from the ED after matching



Source: Mathematica's analysis of Medicare FFS claims

Note: Maryland mean is weighted for observability in Medicare FFS claims. Comparison group mean is weighted for matching and observability.

ED = emergency department.

To test whether the model increased the likelihood of being sent to observation, we estimated impacts of the model, for patients presenting in the ED, on the probability of (1) being admitted, (2) being sent to observation, or (3) being discharged to the community, each shown below (Exhibit B.22). Overall, we do find evidence that the model increased the likelihood of being sent to observation from the ED, almost entirely offset by decreases in the likelihood of being admitted. However, as shown in Exhibit B.21, the likelihood of being sent to observation also increased faster in Maryland than the comparison group during the baseline period (including in our statistical tests which failed to show similar baseline trends between Maryland and the comparison group). This suggests that some, and potentially much, of the diverging trends during the MDAPM and MDTCOC periods could be an extension of baseline trend differences, not true model effects. Additional details on observation stays can be found in Chapter 2.4.1.

Exhibit B.22. Impacts of the Maryland Model on probability of being admitted, sent to observation, or discharged to community from the ED

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Probability of being admitted from ED					
Baseline period (2011–2013)	39.7%	38.3%	1.4pp		
Early MDAPM period					
2014	36.1%	33.8%	2.3pp	0.9** (0.3; 1.4)	2.6%
2015	34.7%	33.2%	1.5pp	0.1 (-0.6; 0.8)	0.3%
2016	34.1%	32.4%	1.7pp	0.3 (-0.7; 1.2)	0.9%
Later MDAPM period					
2017	33.4%	32.7%	0.7pp	-0.7 (-1.6; 0.2)	-2.1%
2018	33.2%	33.6%	-0.4pp	-1.8** (-2.6; -1)	-5.1%
Combined (2017-2018)	33.3%	33.1%	0.2pp	-1.2** (-2.1; -0.4)	-3.5%
MD TCOC period					
2019	32.4%	33.4%	-1.0pp	-2.4** (-3.2; -1.5)	-6.9%
2020	35.9%	37.9%	-2.0pp	-3.4** (-4.4; -2.5)	-8.7%
2021	35.3%	36.8%	-1.5pp	-2.9** (-3.8; -2)	-7.6%
2022	35.2%	36.2%	-1.0pp	-2.4** (-3.3; -1.5)	-6.4%
Combined (2019-2022)	34.6%	36.0%	-1.4pp	-2.8** (-3.6; -1.9)	-7.5%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				-1.1** (-1.8; -0.4)	-3.4pp
2020				-2.2** (-3.1; -1.2)	-5.2pp
2021				-1.7** (-2.6; -0.8)	-4.1pp
2022				-1.2** (-2; -0.4)	-2.9pp
Combined (2019-2022)				-1.5** (-2.3; -0.7)	-4.0pp
Probability of being sent to observation stay from the ED					
Baseline period (2011-2013)	7.0%	6.1%	0.9pp		
Early MDAPM period					
2014	9.4%	8.1%	1.3pp	0.4** (0.1; 0.8)	4.4%
2015	10.0%	8.4%	1.6pp	0.7** (0.3; 1.1)	7.5%
2016	10.4%	8.9%	1.5pp	0.6** (0.2; 0.9)	6.1%
Later MDAPM period					
2017	10.6%	8.7%	1.9pp	1** (0.6; 1.5)	10.4%
2018	11.2%	8.8%	2.4pp	1.5** (1.1; 2)	15.5%
Combined (2017–2018)	10.9%	8.7%	2.2pp	1.3** (0.8; 1.7)	13.5%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
MD TCOC period					
2019	11.7%	8.7%	3.0pp	2.1** (1.6; 2.6)	21.9%
2020	11.9%	7.7%	4.2pp	3.3** (2.8; 3.8)	38.4%
2021	11.3%	7.5%	3.8pp	2.9** (2.5; 3.3)	34.5%
2022	11.0%	7.1%	3.9pp	3** (2.4; 3.5)	37.5%
Combined (2019–2022)	11.5%	7.8%	3.7pp	2.8** (2.4; 3.2)	32.2%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				0.8** (0.5; 1.1)	8.4pp
2020				2** (1.5; 2.6)	24.9pp
2021				1.6** (1.2; 2.1)	21.0pp
2022				1.7** (1.1; 2.3)	24.0pp
Combined (2019–2022)				1.5** (1.1; 1.9)	18.7pp
Probability of being discharged home from ED					
Baseline period (2011–2013)	53.4%	55.7%	-2.3pp		
Early MDAPM period					
2014	54.5%	58.1%	-3.6pp	-1.3** (-1.8; -0.8)	-2.3%
2015	55.3%	58.4%	-3.1pp	-0.8* (-1.6; 0)	-1.4%
2016	55.5%	58.7%	-3.2pp	-0.9 (-1.8; 0.1)	-1.6%
Later MDAPM period					
2017	56.0%	58.6%	-2.6pp	-0.3 (-1.1; 0.5)	-0.5%
2018	55.7%	57.8%	-2.1pp	0.2 (-0.6; 1)	0.4%
Combined (2017–2018)	55.8%	58.1%	-2.3pp	0 (-0.8; 0.8)	-0.0%
MD TCOC period					
2019	55.9%	57.9%	-2.0pp	0.3 (-0.4; 1)	0.5%
2020	52.3%	54.5%	-2.2pp	0.1 (-0.7; 0.9)	0.2%
2021	53.4%	55.7%	-2.3pp	0 (-0.9; 0.9)	-0.0%
2022	53.8%	56.6%	-2.8pp	-0.5 (-1.2; 0.1)	-0.9%
Combined (2019–2022)	54.0%	56.3%	-2.3pp	0 (-0.7; 0.7)	-0.0%
Difference in estimates during MD TCOC period and later MDAPM period					
2019				0.3 (-0.4; 1)	0.5pp
2020				0.1 (-0.7; 1)	0.2pp
2021				0 (-0.8; 0.9)	0.0pp
2022				-0.5 (-1.3; 0.3)	-0.9pp
Combined (2019–2022)				0 (-0.7; 0.7)	0.0pp

* $p < 0.10$; ** $p < 0.05$.
ED=emergency department

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

CI = confidence interval; pp= percentage point.

Given that the model likely reduced the likelihood of being sent to observation, we also wanted to understand how important this mechanism could be, relative to our estimates of how much the model reduced hospital admissions. In other words, could sending more beneficiaries to observation fully explain reductions in admissions we observed from 2014-2022? We estimated this in Exhibit B.23 in several steps. First, we took the estimated model impacts on the probability of being sent to observation (column A), for example, 2.08 percentage points in 2019. We then multiplied these impacts by the total number of ED visits in column B for an estimate of the total number of admissions that were shifted from what would have been an admission to what is now an observation stay in column C (2.08 admissions averted from observation use/100 ED visits * 482,895 ED visits = 10,043 admissions averted from observation use in 2019). Next, we took our estimated impacts on admissions in column D and multiplied by the total number of Medicare FFS beneficiaries in Maryland in column E to get an estimate of the total number of admissions the model prevented in column F (-0.0483 total admissions averted/beneficiary * 802,961 beneficiaries = 38,889 total admissions averted in 2019). Finally, we can compare divide the number of admissions shifted from column C by the total number of admissions prevented in column F for an estimate of the share of the total admissions the model prevents that could be attributable to shifts to observation stays in column G (10,043 admissions averted from observation use / 38,777 total admissions averted = 25.9% in 2019). Our calculations suggest that up to 38% of the total effect on admissions could be explained by shifts to observation stays.

Exhibit B.23. Up to 38% of the Model's effect on total hospital admissions can be explained by shifting beneficiaries from being admitted to being sent to observation

Year	A Impacts on probability of observation stay	B Total ED visits in MD (denominator for A)	C Total admissions shifted (A * B)	D Impacts on admissions (per beneficiary)	E Total Medicare FFS beneficiaries in MD (denominator of D)	F Total admissions prevented (D*E)	G Share of total admissions prevented that could be shifted (C/F)
2014	0.45pp	475,780	2,119	-0.0076	740,351	5,625	37.7%
2015	0.71pp	490,629	3,478	-0.0161	764,546	12,317	28.2%
2016	0.57pp	480,137	2,736	-0.0187	773,262	14,472	18.9%
2017	1.00pp	484,020	4,863	-0.0276	777,570	21,474	22.6%
2018	1.54pp	483,038	7,446	-0.0354	785,850	27,817	26.8%
2019	2.08pp	482,895	10,043	-0.0483	802,961	38,777	25.9%
2020	3.32pp	380,885	12,648	-0.0479	812,664	38,889	32.5%
2021	2.92pp	394,961	11,528	-0.0388	793,244	30,744	37.5%
2022	2.97pp	387,921	11,525	-0.0393	777,372	30,573	37.7%

Source: Mathematica's analysis of Medicare FFS claims data from 2011-2022.

Notes: This table estimates the proportion of total admissions that are explained by shifting beneficiaries to observation stays. We used impacts on the probability of being send to observation from the ED as well as impacts on total admissions shown in Chapter 2.

ED=emergency department, FFS=fee-for-service. pp=percentage points

B.2. Measures, definitions, and file construction for Medicare analyses

This appendix describes how we constructed Medicare FFS claims-based outcomes measures and survey-based patient experience measures, PUMA-level matching variables, and regression covariates for the impact analyses in this report. We first describe in detail how we defined the outcomes measures, starting with the claims-based measures and then the survey-based patient experience measures. For the claims-based measures, we organized this appendix by whether they are measured at the beneficiary-year level or the discharge level (Section B.1). We then describe how we rolled up the claims-based beneficiary- and discharge-level outcomes measures along with the survey measures and other claims- and non-claims-based measures—including beneficiaries' demographic and enrollment characteristics, health status measures, and geographic characteristics—to develop PUMA-level matching variables (Section B.2.1). Finally, we describe the analysis files used for beneficiary- and discharge-level impact models, including definitions of covariates constructed from claims, enrollment, area-level, and patient survey data as well as the files constructed at the hospital-level for analyses of hospital-based patient experience (Section B.2.2).

We constructed annual files with outcomes, matching variables, and regression covariates –for beneficiary-level, discharge-level (episodes), and hospital-level outcomes. The annual claims-based beneficiary file contains one observation per beneficiary per year for all beneficiaries who were observable for at least one month in Medicare FFS claims data during the year (that is, they were alive, enrolled in Medicare Parts A and B FFS, and had Medicare as primary payer). Beneficiaries can be in the file in all years of our analytic period or only one or a limited number of years, depending on their observability status. One annual discharge file contains discharges paid for by FFS Medicare that met denominator inclusion criteria for 30-day unplanned readmissions (Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation 2020) or follow-up after acute exacerbations of chronic conditions (IMPAQ Health 2018). Another annual discharge file contains all Medicare FFS inpatient and outpatient claims involving an ED visit that we used for the ED discharge destination analyses. For analyses of patients' ratings of their personal doctors, the annual file contains survey responses from CAHPS from Medicare FFS and Medicare Advantage respondents in each year along with demographic and enrollment-related characteristics from Medicare administrative files. For analyses of hospital ratings, the analysis file contains hospital-year-level average annual ratings along with covariates that measure characteristics of the hospitals.

B.2.1. Measures and definitions

To construct claims-based outcomes at the beneficiary-year level, we relied on the Medicare FFS Research Identifiable Files (RIFs) claims data from the Virtual Research Data Center. These files provide data on all services funded by Medicare FFS. We used claims data with at least 90 days of runout at the time we pulled the data, the standard for evaluation purposes. We used all claims to measure outcomes, regardless of geography. For example, we included all Medicare claims for a Maryland resident, regardless of whether the beneficiary received the covered services from providers in Maryland or elsewhere. We supplemented these data, as described later, with data from the Medicare Geographic Variation Data Base (GVDB) to measure standardized hospital spending.

B.2.1a. Medicare spending measures

Our measures of Medicare spending include Medicare payments recorded in Parts A and B RIF claims data. For all spending measures, we started by assigning the amount Medicare paid for each service to a year based on the end date (or through date) on the claim. The one exception was for post-acute care claims, for which the services provided can often span many months even if paid in only a single month. In those cases, we apportioned the spending or service use recorded on the claim according to the number of post-acute care days falling in the respective years.

We then summed Part A and Part B payments for the months that a beneficiary was observable in FFS claims that year (that is, the beneficiary was enrolled in Medicare Parts A and B FFS and had Medicare as primary payer) and annualized the payments to account for the number of months the beneficiary was observable in FFS claims. For example, if a beneficiary was observable for 10 of 12 months of the year, and we observed $10,000 in Medicare Parts A and B payments for this beneficiary over 10 months, then their annualized spending would be ($12,000). These amounts exclude the amounts that third parties and beneficiaries paid for deductibles, coinsurance, and copayments. They also exclude Medicare payments for Part D prescription drugs and any Medicare payment amounts on home health interim RAP (request for anticipated payment) claims. We set negative Medicare payments to zero.

Total Medicare FFS spending (dollars per beneficiary per year)

This outcome measures Medicare spending, in dollars per beneficiary per year, for Parts A and B covered services during the year among beneficiaries who were observable for at least one month during the year. It is the sum of Medicare payments across inpatient, outpatient, skilled nursing facility (SNF), home health, hospice, carrier (or Part B), and durable medical equipment claims. This variable excludes nonclaims payments (that is, payments from the CMS to providers that were made separately from claims), though, as we describe in footnote 26, we undid adjustments to the Medicare payment amount on carrier and outpatient claims for ambulatory services that had been adjusted downward because of participation in specific CMS models.

Medicare FFS spending, hospital and non-hospital spending

We also measured Medicare FFS spending for Parts A and B covered services during the year stratified by type of service: hospital and non-hospital spending. Specifically, we constructed the following categories:

1. **Hospital spending** includes spending for Part A inpatient and Part B outpatient claims at short-stay acute care hospitals, critical access hospitals, children’s hospitals, inpatient rehabilitation hospitals, long-term care hospitals, and psychiatric hospitals.

2. **Non-hospital spending** measures the sum of all Parts A and B spending that was not classified as hospital spending according to the earlier definition. Specifically, non-hospital spending is the sum of the following measures:

   2.1. **Post-acute care spending** measures the sum of Part A spending for SNF and home health services, defined as follows:

   2.1.1. **SNF spending** measures all spending for service use recorded in the SNF claims file. It includes spending for SNF services provided in swing beds in short-term acute care hospitals.

   2.1.2. **Home health visit Medicare Part A spending** measures Medicare Part A spending for service use recorded in the home health agency claims file. Medicare Part B also covers home health care, but Part A provides coverage following a qualifying inpatient hospital stay. This measure aims to capture post-acute care home health spending, so we limited spending to home health care claims covered by Part A, including (a small number of) claims covered by Medicare Parts A and B.

   2.2. **Ambulatory care visit with primary care providers and specialist physicians spending** is the sum of the two ambulatory care visit spending measures below.

   2.2.1. **Ambulatory care visit with primary care provider spending** measures Medicare Part B professional (carrier claim) spending for ambulatory visits with primary care practitioners, nurse practitioners, physician assistants, and other advanced practice nurses. It also includes Part B outpatient spending for ambulatory visits at clinics (Federally Qualified Health Centers and Rural Health Clinics).[26]

   2.2.2. **Ambulatory care visit with specialist physicians spending** measures Medicare Part B professional (carrier claim) spending for ambulatory visits with specialist physicians.

   2.3. **Non-hospital Part B drug spending** measures spending for drugs covered by Medicare Part B that is not classified earlier as hospital spending. Specifically, we identified Medicare spending for claims lines in the non-hospital outpatient claims, carrier claims, and durable medical equipment claims files in which the procedure (Healthcare Common Procedural Coding System, or HCPCS) code was for a drug paid for under the average sales price payment system.

[26] Some primary care providers and specialists participating in ACO Reach (previously Global and Professional Direct Contracting), Comprehensive Primary Care Plus, Kidney Care Choices Model, MDPCP, Next Generation ACO, Primary Care First, and the Vermont ACO Model have their Medicare payment amounts on ambulatory claims adjusted downwards because these visits are otherwise covered partially or wholly under capitated arrangements with CMS or covered under a flat visit fee (Primary Care First). We removed these downward adjustments to obtain the amount Medicare would have paid under FFS (in the absence of capitation or the flat visit fee) for these visits. The Primary Care First Model leads to some claims being adjusted upwards of what Medicare FFS would normally pay. We did not undo these adjustments to ensure we captured all Medicare payments.

2.4. **Ambulatory surgical center facility spending** measures facility charges for services at ambulatory surgical centers. Ambulatory surgical center claims were identified by the claim type of service code ("F"). Spending on Part B drugs was excluded (because this spending was captured in the measure described before).

2.5. **Imaging and testing professional spending** measures spending for professional services associated with imaging and testing. Specifically, it includes spending for claim lines in the carrier claims file in which the procedure code was classified as imaging or testing according to the Berenson Eggers Type of Service (BETOS) or Restructured BETOS Classification System (RBCS) algorithm (CMS 2022b) (we applied the RBCS to all claims, but because the RBCS includes Medicare-covered procedure codes starting in 2014 only, we back-filled the imaging and testing variable in the 2011 through 2013 carrier files with any codes that the RBCS did not classify, but that the BETOS algorithm classified as imaging and testing) . Professional spending excludes any outpatient facility charges for imaging and testing conducted in settings for which outpatient facility claims are also submitted.

2.6. **Other non-hospital spending** measures the sum of all Parts A and B spending not captured by any of the measures described before. This measure includes Medicare Part A spending on non-hospital inpatient services[27] and hospice; Part B spending on home health care and ambulatory care visits with behavioral health providers, and Part B spending for non-hospital outpatient, professional (carrier) services, and durable medical equipment not otherwise captured in the measures before (for example, not previously categorized as spending on Part B drugs).

3. **Total Medicare spending plus non-claims payments** measures total spending plus payments made in support of alternative payment models. Specifically, it includes, in Maryland and the comparison group, when applicable, payments for the following programs: Pioneer ACO, ACO Shared Savings Program, Next Generation ACO, Comprehensive Primary Care Initiative, Comprehensive Primary Care Plus, Primary Care First, Global and Professional Direct Contracting (ACO REACH as of 2021), MDPCP, and payments to providers that participated in advanced alternative payment models under the Quality Payment Program. For MDPCP, Comprehensive Primary Care Initiative, and Comprehensive Primary Care Plus, payments comprise all payments, including care management fees, performance-based incentive payments, and comprehensive primary care payments, when applicable. Payments to each of these programs and the total attributed beneficiaries in 2014, 2019, and 2021, are included in Exhibit B.24.

[27] This category includes claims from facilities that are excluded from our definition of hospital spending, such as religious non-medical health care institutions.

Exhibit B.24. Average non-claims payments to CMS and total number of participating beneficiaries for select programs, 2014, 2019, and 2021

	2014		2019		2021	
	Maryland	Comparison	Maryland	Comparison	Maryland	Comparison
Average payment per beneficiary, $						
Pioneer ACO	$0.04	$2.35	n.a.	n.a.	n.a.	n.a.
Medicare Shared Savings Program	$12.21	$9.59	$12.58	$43.34	$21.09	$60.32
Next Gen ACO	n.a.	n.a.	n.a.[a]	$12.99	n.a.[a]	$8.39
General and Professional Direct Contracting (ACO REACH)	n.a.	n.a.	n.a.	n.a.	$0.49	$1.46
Comprehensive Primary Care	n.a.[a]	$1.89	n.a.	n.a.	n.a.	n.a.
Comprehensive Primary Care Plus	n.a.	n.a.	n.a.[a]	$15.32	n.a.[a]	$14.80
Maryland Primary Care Program	n.a.	n.a.	$75.86	n.a.[a]	$194.72	n.a.[a]
Primary Care First	n.a.	n.a.	n.a.	n.a.	n.a.[a]	$5.89
Advanced Alternative Payment Model payments	n.a.	n.a.	$8.86	$7.82	$27.11	$19.57
Number of participating beneficiaries						
Pioneer ACO	161	143,101				
Medicare Shared Savings Program	153,781		134,481	3,054,583	74,271	2,820,722
Next Gen ACO			206	262,621	186	195,382
General and Professional Direct Contracting (ACO REACH)					157	93,584
Comprehensive Primary Care	65					
Comprehensive Primary Care Plus			880	530,132	809	511,110
Maryland Primary Care Program			210,371	1,893	397,349	3,145
Primary Care First					1,107	183,386
Advanced Alternative Payment Model payments			751,129	7,161,991	737,204	7,131,042

Source: Mathematica's analysis of Medicare claims data and the Master Data Management file.

Note: This table shows per-capita non-claims payments in 2014, 2019, and 2021, as well as the total number of participating beneficiaries in each program. The analysis is based on beneficiary attribution in the Master Data Management file maintained by CMS to track participation in CMS initiatives. Information on payments comes from Mathematica's analysis of implementation contractor data for the Maryland Primary Care Program, Primary Care First, Comprehensive Primary Care Plus and Comprehensive Primary Care, publicly available sources on total share savings payments for ACO programs, and CMS data on payments to individual providers for Advanced Alternative Payment model payments. Per-capita averages include all beneficiaries in Maryland and the comparison group (includes those with zero payments).

[a] A small number of beneficiaries are assigned to programs not officially active in that area. For example, there are a small number of Comprehensive Primary Care Plus–assigned beneficiaries in Maryland in 2019. This is because we use beneficiary address at the start of the year to determine location for individual beneficiaries, not the practice location. If a beneficiary crosses state lines to receive care, or moves within the year, they could show up as participating in a location without any official participation. We include these beneficiaries in our analysis, but they represent negligible non-claims dollar amounts (< $1) that are unlikely to influence our results.

ACO = accountable care organization; n.a. = not applicable.

B.2.1b. Service use measures

Intensity of hospital care (measured by standardized hospital spending)

We computed measures of annualized standardized hospital spending using the Medicare GVDB, produced by the CMS Office of Information Products and Data Analytics. The database includes claim-level standardized payment amounts for Part A claims (inpatient, SNF, hospice, and home health) and Part B institutional (outpatient) claims. We merged the standardized payment amounts onto the RIF files (at the claim level for Part A claims and Part B institutional claims). Then we calculated standardized hospital payments across the same set of claims in the hospital spending category described above with the standardized payment amounts from the GVDB in place of actual hospital payment amounts. Standardized spending removes differences in spending across claims because of difference in the prices paid to different providers (for example, those from wage indices in different parts of the country or Health Services Cost Review Commission rate setting), so it measures intensity of service use in aggregate. For this progress report, for all years (2012 to 2020), we used the GVDB 24-month files. For 2021, we used the GVDB 18-month file, which contains less runout than the 24-month files but was the only version available at the time of these analyses.

All-cause acute care hospital admissions (number of admissions per beneficiary per year)

This measure is the annualized number of hospitalizations for short-stay acute hospitals, critical access hospitals, and children's hospital admissions reported in the RIF inpatient claims file for the beneficiary during the year. Multiple claims for acute admissions that involved transfers between hospitals were combined into a single record, as were multiple claims for the same beneficiary at the same facility with overlapping dates, so these count as one admission. We excluded hospitalizations for psychiatric care, inpatient rehabilitation stays, and long-term hospital stays.

Outpatient ED visits and observation stays (number of visits per beneficiary per year)

This measure is the annualized number of outpatient ED visits and observation stays for the beneficiary during the year that do not lead to a hospitalization. Visits that do not lead to a hospitalization are identified in the outpatient department RIF hospital claims file using revenue center line items equal to 045X or 0981 (emergency room care), 0762 (treatment or observation room), or 0760 (treatment or observation room—general classification). We counted a visit as an observation stay if it was longer than eight hours and had a corresponding HCPCS code of G0378 (hospital observation services per hour). We then capped the number of either type of visit (observation stays and ED visits) to one per day.

B.2.1c. Quality of care measures

Potentially preventable admissions (number of admissions per beneficiary per year)

This measure is the annualized number of hospitalizations for short-stay acute hospital, critical access hospital, and children's hospital admissions reported in the inpatient claims file for the beneficiary during the year in which the admission met the criteria for the Prevention Quality Indicators (PQI) overall composite measure (PQI #90). To construct this measure, we applied the Agency for Healthcare Research and Quality's 2020 Quality Indicators Software to all inpatient hospital claims for acute stays (defined earlier) and then counted the number of hospital admissions for the beneficiary each year that the software flagged as being admissions for one of the following PQIs: diabetes short-term complications (PQI #01), diabetes long-term complications (PQI #03), chronic obstructive pulmonary disease or asthma in older adults (PQI #05), hypertension (PQI #07), heart failure (PQI #08), community-acquired pneumonia (PQI #11), urinary tract infection (PQI #12), uncontrolled diabetes (PQI #14), asthma in younger adults (PQI #15), or lower-extremity amputation among patients with diabetes (PQI #16) (AHRQ n.d.[a]).

B.2.1d. Population health measures

Use of Medicare Diabetes Prevention Program Services (yes or no for the beneficiary during the year)

This measures whether the beneficiary received any Medicare Diabetes Prevention Program services during the year (yes or no). A beneficiary was considered to have received Medicare Diabetes Prevention Program services if they had at least one outpatient or carrier claim with procedure code 0403T, 0488T, G9873, G9874, G9875, G9876, G9877, G9878, G9879, G9882, G9883, G9884, G9885, G9880, G9881, G9890, or G9891. Medicare started funding Medicare Diabetes Prevention Program services in 2018. Therefore, this outcome will have a value of 0 for all beneficiaries from 2011 to 2017.

B.2.1e. Quality of care outcomes measured at the discharge-year level

30-day post-discharge unplanned readmission (yes or no for the event)

We used Medicare FFS RIF inpatient claims and enrollment data for this measure. The analytic file has one observation for each inpatient discharge. Beneficiaries can be included in the file once, more than once, or not at all depending on how many discharges they had. Multiple claims for acute admissions that involved transfers between hospitals were combined into a single record, as were multiple claims for the same beneficiary at the same facility with overlapping dates, so these count as one discharge.

The all-cause 30-day post-discharge unplanned readmission measure indicates whether the discharge (the index admission) was followed by an unplanned hospital admission within 30 days. An unplanned readmission is defined as any hospitalization that does not follow an established plan of care (examples of planned admissions include those for chemotherapy and planned admission for transplant surgery). The measure equals 1 if there was an unplanned readmission within 30 days of discharge to any hospital, regardless of whether the readmission occurred at the same hospital or a different hospital. The measure equals 0 if there was no unplanned readmission within 30 days.

Our definition of this measure is based on the Yale readmission measure developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2020) used in the Hospital Readmission Reduction Program under Section 3025 of the Affordable Care Act. An admission that counts as a readmission because it fell within 30 days of an earlier index stay can also count as an index stay for a potential subsequent readmission as long as it meets the index admission inclusion criteria. We count an index admission in a year if the discharge date is in that year. We then look for an unplanned readmission within 30-days of that index admission (the readmission could occur in the following year)

Timely follow-up after acute exacerbations of chronic conditions (yes or no for the event)

This measures whether follow-up was received within the time frame recommended by clinical practice guidelines in a non-emergency outpatient setting following an ED visit or hospitalization for one of the following six chronic conditions: hypertension, asthma, heart failure, coronary artery disease, chronic obstructive pulmonary disease, or diabetes mellitus (Type I or Type II). IMPAQ Health (2018) developed the measure specifications. The Health Services Cost Review Commission has included improvement on this measure as one of its quality goals in Statewide Integrated Health Improvement Strategy.

To develop this measure, we first identified hospital admissions and outpatient emergency visits and observation stays that met the denominator criteria for one of the six chronic conditions. Unlike the readmission measure defined before, this measure is not strictly at the inpatient discharge level; the denominator includes outpatient ED visits and observation stays as well as inpatient discharges. Nonetheless, we group the measure with other discharge-level outcome measures because we analyzed the outcome with the same methods. We then applied the measure's additional denominator inclusion criteria with just one minor modification (that is, we included index events in December because we had claims data for the subsequent year). We then flagged qualifying events with timely follow-up—an outpatient or carrier claim for the same patient after the index event for a non-emergency outpatient visit that constitutes appropriate follow-up (for example, a general office visit) using the IMPAQ (2018) code set (Because the IMPAQ code set has not been updated in several years, it is possible that we underestimated follow-up in recent years. However, we expect that we likely underestimated follow-up similarly for Maryland and comparison group). The follow-up visit must occur within the condition-specific time frame to be considered timely: within 7 days of the date of discharge for hypertension; within 14 days for asthma, heart failure, and coronary artery disease; and within 30 days for chronic obstructive pulmonary disease and diabetes.

B.2.1f. ED discharge destination

We used Medicare FFS RIF inpatient and outpatient claims and enrollment data for this measure. The analytic file contains one observation for each ED visit, which could be an inpatient claim because the beneficiary was admitted to the hospital from the ED or an outpatient claim because the beneficiary was not admitted or the facility was allowed to bill separately for ED and inpatient care. Beneficiaries can be included in the file once, more than once, or not at all depending on how many ED visits they had. If an ED visit involved multiple claims, we limited the analysis file to one claim for that ED visit. For example, if a beneficiary had multiple outpatient ED claims from different facilities on the same day, we considered this to be one ED visit. Similarly, if a beneficiary had an outpatient ED claim from one facility and an inpatient claim whose from date equaled the thru date on the outpatient ED claim or one day later, and the inpatient claim indicated that the beneficiary came through the ED of the (second) hospital, we treated this set of claims as one ED visit.

The denominator includes all ED visits, except for ED visits (1) that ended in death (PTNT_DSCHRG_STUS_CD = 20 on an outpatient ED claim); (2) in which the beneficiary left against medical advice (PTNT_DSCHRG_STUS_CD = 07 on an outpatient ED claim); (3) with a principal diagnosis for a psychiatric condition;[28] or (4) associated with an elective hospitalization (CLM_IP_ADMSN_TYPE_CD = 3), for which we assumed the ED visit covered planned, pre-admission services within the three days before admission.

We defined three mutually exclusive numerator events:

1. **Admitted as an inpatient.** We flagged inpatient claims for non-elective admissions as meeting the numerator criteria if the claim had an ED revenue center code (as defined above). We flagged outpatient ED claims as resulting in admission if we found an inpatient claim for a non-elective admission for the same beneficiary with a from date equal to the thru date of the ED visit or one day later. ED visits flagged as being admitted might have also involved an observation stay.
2. **Sent to observation.** We flagged outpatient ED claims as meeting this numerator criteria if there was no evidence that the beneficiary was admitted to a hospital and if the ED visit claim indicated that the visit also involved an observation stay (as defined above) or if we found a separate outpatient claim for an observation stay for the same beneficiary with a from date that equaled the thru date of the ED visit or one day later.
3. **Discharged to the community.** All outpatient ED visit claims that were not categorized as having led to an admission or sent to observation were flagged as discharged to the community.

[28] We exclude psychiatric ED claims because the set of inpatient claims that we use to measure acute, all-cause hospitalizations also excludes claims with a principal psychiatric diagnosis, and we would be unable to determine whether a psychiatric-related ED visit resulted in an inpatient admission.

B.2.1g. Quality of care measures from patient experience surveys

Patients' rating of their personal doctor (score from 0 to 100)

We used the FFS CAHPS and Medicare Advantage CAHPS RIFs from the Virtual Research Data Center to construct survey respondent-level files for Maryland and the comparison group. The FFS and Medicare Advantage CAHPS files were linked to the Medicare beneficiary analytic files with the annual claims-based outcomes using each beneficiary's unique beneficiary identifier. We limited the CAHPS data to respondents who received a non-zero or non-missing survey weight. The file has one observation per respondent, grouped by year.

This CAHPS questionnaire asks respondents to rate their personal doctor. The rating question states: "Using any number from 0 to 10, where 0 is the worst personal doctor possible and 10 is the best personal doctor possible, what number would you use to rate your personal doctor?" Therefore, the measure that rates beneficiaries' personal doctor includes all responses to this question, and the measure of beneficiaries' primary care provider is restricted to those who answer that their personal doctor is not a specialist. We then multiplied this measure by 10 to put in on a scale of 0 to 100. Responses to this question are heavily top-coded, with means in Maryland and the comparison group above 90 in most years (Appendix B.1.5).

Although we limited our claims-based analyses to FFS enrollees because of data availability, we included Medicare Advantage enrollees in these analyses for several reasons: (1) having Medicare Advantage enrollees in the analysis sample reduces concerns that impact estimates could be biased because of differential enrollment in Medicare Advantage over time among beneficiaries with different health care needs and expected spending (see section B.1.9); (2) doctors in Maryland participating in MDPCP (one key mechanism for improving patient experience scores) are scored and incentivized (via value-based payments) based on scores from FFS and Medicare Advantage CAHPS; (3) we improve the reliability and power of these survey-based analyses by including more survey respondents in these analyses.

In processing the data, we noticed a data anomaly in the years 2015 and 2016. Specifically, survey response rates and mean ratings dropped considerably in those years for Maryland and the comparison group. We also observed that the CAHPS survey weights accounted for this drop, and when we applied the survey weights, we did not see large drops in the number of survey respondents (Appendix B.1.5). In testing our regressions, we did not see a material difference in the results with or without including 2015 and 2016. As such, we decided to continue to use these years in our primary regression models.

Patients' rating of their hospital

To assess patients' rating of their hospital, we used information from the Hospital CAHPS survey contained as part of the publicly available Hospital Compare database in each year from 2011 to 2022 (CMS 2021b). Reporting for the Hospital CAHPS survey was suspended in the first half of 2020 because of COVID-19. Values from 2020 files are based on surveys from the second half of 2020 only.

The public-use files are based on survey responses from patients who had an inpatient hospitalization during the year, administered to patients between 48 hours and 6 weeks after discharge from the hospital. Importantly, scores contained in the public-use Hospital CAHPS files are averages based on Medicare and non-Medicare patients. Specifically, the survey asks, "Using any number from 0 to 10, where 0 is the worst hospital possible and 10 is the best hospital possible, what number would you use to rate this hospital during your stay?" Survey responses are then averaged for each individual hospital in the year and then reported in the public-use data files at the hospital-year level. Before being publicly reported, data are adjusted for the effects of patient-mix and mode of survey administration (HCAHPS, 2022). Because the mean score was not reported in all years of our analysis, we used the percentage of patients who rated their hospital a 9 or 10 out of 10 (which is available in all years) instead.

Several major hospital systems in Maryland did not report Hospital CAHPS scores in 2016. Because of the influence these hospitals have on the Maryland mean, we chose to exclude the year 2016 from all of our analyses of this measure.

B.2.2. Matching and analytic files construction

B.2.2a. PUMA-year-level file with variables for developing the matched comparison group

To develop the PUMA-year-level matching file, we first assigned each beneficiary to the PUMA associated with the beneficiary's mailing address zip code in each year. We then rolled-up the beneficiary-year-level demographic and enrollment file and claims-based outcomes file to the PUMA and year level—that is, one observation per PUMA per year—and calculated the mean value of each variable over all Medicare FFS beneficiaries who resided in that PUMA in that year, weighted by the number of months that each beneficiary was observable in Medicare claims in that year. In addition, we linked the CAHPS patient experience data to beneficiaries in the beneficiary-year-level files and rolled up the survey data to the PUMA-year level based on beneficiaries' assigned PUMAs. We also rolled up the discharge-year file to the PUMA and year level based on the beneficiary's home PUMA (even if the beneficiary was hospitalized outside the PUMA) and calculated the mean value of discharge-related outcomes over all discharges among beneficiaries in each PUMA. We similarly calculated matching variables or variables for checking balance between Maryland and comparison PUMAs from beneficiary-level and discharge-level claims measures, such as COVID-19-related hospitalizations and outpatient ED visits and observation stays and a PUMA-level measure of hospital market concentration and rolled these up to the PUMA-year-level.

As part of the process for constructing the matching files, we also rolled up data from other sources that we merged to the beneficiary-year-level file, including (1) American Community Survey (ACS) data for characteristics of the beneficiaries' zip codes; (2) Health Resources Services Administration data for the primary care shortage area score of each zip code; and (3) HCC scores and individual condition categories from beneficiary-year-level tables on the Virtual Research Data Center. Similarly, we also merged data from Hospital Compare and the Inpatient Prospective Payment System Historical Impact Files to the discharge-year-level file by hospital that were then rolled-up to the PUMA level. Finally, we obtained survey data from the CDC's Diabetes Atlas, which is derived from respondent-level data from the Behavioral Risk Factor Surveillance System. The Diabetes Atlas data are based on the Behavioral Risk Factor Surveillance System data and provide annual estimates of county-level age-adjusted obesity prevalence, diabetes incidence, and diabetes prevalence for adults older than 20 in the county. We mapped counties to their corresponding PUMAs and, for PUMAs with more than one county, we derived PUMA-level estimates using a weighted average based on county population size using the ACS data.

Exhibit B.25 describes the variables included in the matching algorithm or in balance checks. All demographic, enrollment, and geographic variables reflect the characteristics of the PUMAs in 2013. The claims-based outcomes measures and some survey measures include variables for both levels of the outcomes in 2013 and trends in the mean yearly rate of change over the full baseline period, 2011 to 2013.

Exhibit B.25. PUMA-level matching variables

Variables	Data source	Definition
Medicare FFS spending		
Medicare FFS Part A post-acute care spending: baseline levels	Medicare FFS claims	Mean Part A post-acute care spending (that is, for SNF and home health care covered under Part A) per beneficiary per year, calculated over all Medicare FFS beneficiaries in the PUMA in 2013
Medicare FFS Part A post-acute care spending: baseline trends	Medicare FFS claims	Mean change in Part A post-acute care spending per beneficiary per year, 2012 to 2013
Medicare FFS non-hospital spending: baseline levels	Medicare FFS claims	Mean non-hospital spending (that is, all Medicare spending for services provided outside of acute care hospitals, excluding Part D drugs) per beneficiary per year, calculated over all Medicare FFS beneficiaries in the PUMA in 2013.
Medicare FFS non-hospital spending: baseline trends	Medicare FFS claims	Mean change in non-hospital spending per beneficiary per year, 2011 to 2013
Standardized Medicare FFS spending		
Standardized hospital spending: baseline levels	Medicare FFS claims (GVDB files)	Mean standardized hospital spending per beneficiary per year, calculated over all FFS beneficiaries in the PUMA in 2013
Standardized hospital spending: baseline trends	Medicare FFS claims (GVDB files)	Mean change in standardized hospital spending per beneficiary per year, 2011 to 2013
Service use		
All-cause acute care hospitalizations: baseline levels	Medicare FFS claims	Mean number of all-cause acute care hospitalizations per 1,000 beneficiaries per year, calculated over all Medicare FFS beneficiaries in the PUMA in 2013
All-cause acute care hospitalizations: baseline trends	Medicare FFS claims	Mean change in all-cause acute care hospitalizations per 1,000 beneficiaries per year, 2011 to 2013

Variables	Data source	Definition
Outpatient ED visit and observation stays: baseline levels	Medicare FFS claims	Mean number of outpatient ED visits and observation stays per 1,000 beneficiaries per year, calculated over all Medicare FFS beneficiaries in the PUMA in 2013
Outpatient ED visits and observation stays: baseline trends	Medicare FFS claims	Mean change in outpatient ED visits and observation stays per 1,000 beneficiaries per year, 2011 to 2013
Quality		
Potentially preventable hospitalizations: baseline levels	Medicare FFS claims	Mean number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all Medicare FFS beneficiaries in the PUMA in 2013
Potentially preventable hospitalizations: baseline trends	Medicare FFS claims	Mean change in the number of potentially preventable hospitalizations per 1,000 beneficiaries per year, 2011 to 2013
30-day post-discharge unplanned readmissions: baseline levels	Medicare FFS claims	Percentage of hospitalizations that met the criteria for an index stay and were followed by a 30-day all-cause unplanned readmission, calculated over all index stays for Medicare FFS beneficiaries in the PUMA in 2013
30-day post-discharge unplanned readmissions: baseline trends	Medicare FFS claims	Mean change in the percentage of index hospitalizations followed by a 30-day all-cause unplanned readmission, 2011 to 2013
Timely follow-up after a discharge for acute exacerbations of chronic conditions: baseline levels	Medicare FFS claims	Percentage of all hospitalizations or outpatient ED visits or observation stays for any of six chronic conditions (hypertension, asthma, heart failure, coronary artery disease, chronic obstructive pulmonary disease, or diabetes mellitus) that had a non-emergency outpatient follow-up visit within the relevant time frame calculated over all hospitalizations for Medicare FFS beneficiaries in the PUMA in 2013
Timely follow-up after a discharge for acute exacerbations of chronic conditions: baseline trends	Medicare FFS claims	Mean change in the percentage of hospitalizations or outpatient ED visits or observation stays for the six chronic conditions that had a non-emergency outpatient follow-up visit within the relevant time frame, 2011 to 2013
Patients' ratings of their hospital care: baseline levels	Hospital Compare	Percentage of all Medicare FFS discharges from hospitals that had a rating of 9 or 10 out of 10 in Hospital Compare, calculated over all hospitalizations for Medicare FFS beneficiaries in the PUMA in 2013
Patients' ratings of their hospital care: baseline trends	Hospital Compare	Mean change in the percentage of all Medicare FFS discharges from hospitals that had a rating of 9 or 10 out of 10 in Hospital Compare, 2011 to 2013
Beneficiaries' ratings of their PCP: baseline levels	CAHPS	Mean provider rating among all Medicare FFS and MA beneficiaries in the PUMA whose personal doctor is not a specialist and who responded to the CAHPS survey, 2013

Variables	Data source	Definition
Population health		
Obesity prevalence	CDC's Diabetes Atlas[a]	Mean age-adjusted obesity rates for all residents ages 20 and older in the PUMA.
Medicare beneficiaries' characteristics in 2013		
Age	Medicare enrollment data	Mean age of all FFS beneficiaries in the PUMA
Sex	Medicare enrollment data	Percentage of all FFS beneficiaries in the PUMA who are female
Race and ethnicity		
Black	Medicare enrollment data	Percentage of FFS beneficiaries in the PUMA who are Black
Non-Hispanic White	Medicare enrollment data	Percentage of FFS beneficiaries in the PUMA who are non-Hispanic White
Rural residence, (3 variables)	Medicare enrollment data (zip code) and Census Urban and Rural classification by ZCTA (rural)	(1) Average percentage of the population living in a rural area in the PUMA; (2) categorical variable for quartiles of the average percentage of the population living in a rural area in the PUMA
Disabled or ESRD	Medicare enrollment data	Percentage of FFS beneficiaries in the PUMA with original reason for entitlement of disability and/or ESRD
HCC score	Medicare HCC files	Mean HCC score for all FFS beneficiaries in the PUMA
Diabetes for FFS beneficiaries	Medicare HCC files	Percentage of FFS beneficiaries in the PUMA with diabetes
Diabetes among all adults	BRFSS	Percentage of adult population in the PUMA with diabetes
Population characteristics of zip code (2011–2013)		
Percentage Hispanic	ACS	Mean percentage of Hispanic residents, calculated across each Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percentage below federal poverty level, adjusted for cost of living	ACS Census supplemental poverty measure	Mean percentage of residents living below the federal poverty level, calculated across each Medicare beneficiary's zip code for all beneficiaries in the PUMA, and adjusted for cost of living using the Census supplemental poverty measure
Percentage living in multi-unit structure, mobile home or group quarters	ACS	Mean percentage of residents living in multi-unit structures, mobile homes, or group quarters, calculated across each Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percentage older than 65	ACS	Mean percentage of residents ages 65 years and older, calculated across each Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percent younger than 18	ACS	Mean percentage of residents ages birth to 18 years, calculated across each Medicare beneficiary's zip code for all beneficiaries in the PUMA

Variables	Data source	Definition
Percentage with a high school degree or equivalent	ACS	Mean percentage of residents with a high school degree or equivalent, calculated across every Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percentage who speaks English well	ACS	Mean percentage of residents who speak English well, calculated across every Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percentage living in crowded home	ACS	Mean percentage of residents living in a crowded home, calculated across every Medicare beneficiary's zip code for all beneficiaries in the PUMA
Percentage without a vehicle	ACS	Mean percentage of residents without a vehicle, calculated across every Medicare beneficiary's zip code for all beneficiaries in the PUMA
Characteristics of the health care system in the PUMA in 2013		
Medicare coverage	Medicare enrollment data and ACS	Percentage of PUMA residents enrolled in Medicare
At least one acute care hospital, yes or no	Medicare FFS claims	= 1 if the PUMA has one or more acute care hospitals = 0 if the PUMA has no acute care hospital
Number of hospital beds	IPPS	A count of the number of hospital beds in the PUMA
Discharges from a major teaching hospital	Medicare FFS claims and Hospital Compare	Percentage of all hospitalizations from major teaching or very major teaching academic medical center (defined as according to a resident to bed ratio of greater than 0.25), calculated over all hospitalizations for Medicare FFS beneficiaries in the PUMA
Herfindahl-Hirschman Index	Medicare FFS claims	This variable measures the relative amount of competition in the market.[b] For each hospital in each PUMA, we calculated its market share as the percentage of discharges in the PUMA from that hospital. We then squared the market share value of each hospital and summed the squared values across all hospitals in a PUMA. Markets with higher summed values have less competition (and more market concentration) relative to markets with lower summed values (and less market concentration).
PCPs in practices (TINs) that hare small	Medicare FFS claims	Percentage of PCPs who practice in small practices in the PUMA (small is defined as 1 NPI per TIN, or a solo practice)
PCPs in practices (TINs) that are large	Medicare FFS claims	Percentage of PCPs who practice in large practices in the PUMA (large is defined as 6 or more NPIs per TIN)
Number of PCPs per 1,000 Medicare beneficiaries	Medicare FFS claims and enrollment data	For each PUMA, this is the total number of PCPs practicing in the PUMA divided by the total number of Medicare beneficiaries (FFS and Medicare Advantage) and multiplied by 1,000

Variables	Data source	Definition
Health professional primary care shortage area score	HRSA	Mean value of the health professional primary care shortage area score (higher indicates greater shortage) associated with the zip code of all FFS beneficiaries in the PUMA
Subgroups		
Characteristics of Black beneficiaries, baseline		
HCC score, mean	Medicare HCC files	Mean HCC score among Black beneficiaries in the PUMA in 2013
Potentially preventable admissions: baseline levels	Medicare FFS claims	Mean number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all Black Medicare FFS beneficiaries in the PUMA in 2013
Potentially preventable admissions: baseline trends	Medicare FFS claims	Mean change in the number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all Black Medicare FFS beneficiaries in the PUMA, 2011 to 2013
30-day post-discharge unplanned readmissions: baseline levels	Medicare FFS claims	Percentage of hospitalizations that met the criteria for an index stay and were followed by a 30-day all-cause unplanned readmission, calculated over all index stays for Black Medicare FFS beneficiaries in the PUMA in 2013
30-day post-discharge unplanned readmissions: baseline trends	Medicare FFS claims	Mean change in the percentage of index hospitalizations followed by a 30-day all-cause unplanned readmission for Black Medicare FFS beneficiaries in the PUMA, 2011 to 2013
Characteristics of Non-Hispanic White beneficiaries, baseline		
HCC score, mean	Medicare HCC files	Mean HCC score among all non-Hispanic White beneficiaries in the PUMA in 2013
Potentially preventable admissions: baseline levels	Medicare FFS claims	Mean number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all non-Hispanic White Medicare FFS beneficiaries in the PUMA in 2013
Potentially preventable admissions: baseline trends	Medicare FFS claims	Mean change in the number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all non-Hispanic White Medicare FFS beneficiaries in the PUMA, 2011 to 2013
30-day post-discharge unplanned readmissions: baseline levels	Medicare FFS claims	Percentage of hospitalizations that met the criteria for an index stay and were followed by a 30-day all-cause unplanned readmission, calculated over all index stays for non-Hispanic White Medicare FFS beneficiaries in the PUMA in 2013
30-day post-discharge unplanned readmissions: baseline trends	Medicare FFS claims	Mean change in the percentage of index hospitalizations followed by a 30-day all-cause unplanned readmission for non-Hispanic White Medicare FFS beneficiaries in the PUMA, 2011 to 2013

Variables	Data source	Definition
Number of potentially preventable admissions, 2013	Medicare FFS claims	Mean number of potentially preventable hospitalizations per 1,000 beneficiaries per year, calculated over all non-Hispanic White Medicare FFS beneficiaries in the PUMA in 2013
COVID-19 checks, 2019–2021		
COVID-19 inpatient admissions	Medicare FFS claims	Mean number of COVID-19 inpatient visits per 1,000 beneficiaries per year, calculated across all Medicare FFS beneficiaries in the PUMA in 2020 and 2021
COVID-19 outpatient ED visits and observation stays	Medicare FFS claims	Mean number of COVID-19 outpatient ED visits or observation stays per 1,000 beneficiaries per year, calculated across all Medicare FFS beneficiaries in the PUMA in 2020 and 2021
Excess outpatient ED visits and observation stays (2020 minus 2019)	Medicare FFS claims	The difference between the mean number of outpatient ED visits and observation stays in 2020 minus the mean number in 2019 for each PUMA
Excess ED visits and observation stays ending in an inpatient stay (2020 minus 2019)	Medicare FFS claims	The difference between the mean number of ED visits and observation stays that ended in an inpatient stay in 2020 minus the mean number in 2019 for each PUMA
Excess all-cause acute care hospital admissions (2020 minus 2019)	Medicare FFS claims	The difference between the mean number of all-cause acute care hospitalizations in 2020 minus the mean number in 2019 for each PUMA
Excess surgical hospitalizations (2020 minus 2019)	Medicare FFS claims	The difference between the mean number of surgical hospitalizations in 2020 minus the mean number in 2019 for each PUMA
Excess elective hospitalizations (2020 minus 2019)	Medicare FFS claims	The difference between the mean number of elective hospitalizations in 2020 minus the mean number in 2019 for each PUMA

[a] See this page for CDC's Diabetes Atlas: https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html.

[b] For an overview of the Herfindahl-Hirschman Index, see this page from the U.S. Department of Justice: https://www.justice.gov/atr/herfindahl-hirschman-index. Accessed September 29, 2021.

BRFSS = Behavioral Risk Factor Surveillance System; CAHPS = Consumer Assessment of Healthcare Providers and Systems; CDC = Centers for Disease Control and Prevention; ED = emergency department; ESRD = end-stage renal disease; FFS = fee for service; HCC = Hierarchical Condition Category; HRSA = Health Resources and Services Administration; IPPS = Inpatient Prospective Payment System; MA = Medicare Advantage; NPI = National Provider Identifier; PCP = primary care physician; PUMA = Public Use Microdata Area; SNF = skilled nursing facility; TIN = Taxpayer Identification Number; ZCTA = Zip Code Tabulation Area.

B.2.2b. Beneficiary-year and discharge-year level files for impact analyses

We fit regression models at the beneficiary-year or discharge-year level, as relevant, for the claims-based outcomes and for the CAHPS-based patient experience measures to assess impacts of the Maryland Model on key outcomes. The regression models use the same beneficiary-year- and discharge-year-level files described in Section B.2.1 as inputs to the PUMA-year files. Briefly, the beneficiary-year-level analytic file contains one observation PBPY for all beneficiaries ever enrolled in Medicare. For the analyses of claims-based measures, we then limited the file to those who were observable for at least one month in Medicare FFS claims data during the year (that is, they were enrolled in Medicare Parts A and B FFS and had Medicare as primary payer). For the analyses of CAHPS data, we included all respondents to the survey who were enrolled in FFS and met the criteria described above or who were enrolled in Medicare Advantage with Medicare as primary payer (by definition, Medicare Advantage enrollees are enrolled in Parts A and B). Beneficiaries can be in the file in all years of our analytic period or only one or a limited number of years, depending on their observability status.

The construction of this file involved the following steps:

1. Pulling enrollment and demographic information for the full Medicare population (that is, all beneficiaries who were ever enrolled in Medicare) during each year from the Master Beneficiary Summary File and Enrollment Database
2. Identifying the first FFS observable month, if any, among the full Medicare population (many beneficiaries are never observable during the year because, for example, they are enrolled in Medicare Advantage for the full year or have employer-sponsored insurance as primary payer or only have Part A or only Part B coverage during enrolled months) for the claims-based analyses
3. Constructing variables to reflect demographic and enrollment characteristics for the year, as described in Table B.2; characteristics that could change within any year, such as dual eligibility status and residence (based on zip code); and those that are characterized based on beneficiaries' data in the first observable month of the year
4. Developing claims-based measures for all observable months for all FFS beneficiaries and merging these measures to the beneficiary-level file by unique beneficiary identifier
5. Obtaining patient survey measures for all FFS and Medicare Advantage enrollees who responded to the CAHPS in each year
6. Annualizing the claims-based measures based on the number of months observable (except for binary variables, such as "any hospitalization")
7. Merging on data from external sources (this included merging HCC scores and Master Beneficiary Summary File chronic condition categories by unique beneficiary identifier and year) and merging on characteristics of beneficiaries' PUMA from the ACS by PUMA and year (we used the ACS five-year files, which combine data for each PUMA across five years, so each PUMA will have the same values of the ACS variables across the five-year period covered)

8. Applying a final set of exclusion criteria for each year's file to exclude beneficiaries from the analytic sample if we could not map them to a location in the United States (either because they lived outside the United States or had bad zip code data)

The discharge-year-level files contain one observation per Medicare FFS discharge or ED visit per year. For the readmissions and follow-up after exacerbations of acute conditions measures, the file contains either a discharge from an acute inpatient hospital, regardless of the reason for the hospitalization, or an outpatient discharge from the ED or observation unit with a diagnosis code for any of the six chronic conditions included in the follow-up after acute exacerbations of chronic conditions measure. For the ED access measure, the file contains one record per ED visit. Each record in this file represents a single discharge or ED visit, meaning any hospital use involving more than one claim was collapsed into a single record (for readmissions and the follow-up measure, we retained inpatient diagnosis and procedure codes from the first and last claims in the stay). For the follow-up measure, we limited outpatient ED visits and observation stays to one per day; if any outpatient ED or observation claim on the same day contained relevant diagnoses for the six chronic conditions, we included it in the discharge-level file. Similarly, for the ED access measure, we limited ED visits to one per day.

We then merged demographic and enrollment characteristics, and Medicare CCW condition categories to use as covariates in the regressions onto the files by unique beneficiary identifier and year. CCW conditions included are the following original conditions: acquired hypothyroidism; acute myocardial infarction; Alzheimer's disease and related disorders or senile dementia; anemia; asthma; atrial fibrillation; benign prostatic hyperplasia; cancer – breast; cancer – colorectal; cancer – endometrial; cancer – lung; cancer – prostate; cataract; chronic kidney disease; chronic obstructive pulmonary disease and bronchiectasis; glaucoma; heart failure; hip/pelvic fracture; hyperlipidemia; hypertension; ischemic heart disease; osteoporosis; rheumatoid arthritis/osteoarthritis; and stroke/transient ischemic attack. We also included the following other chronic and potentially disabling conditions: blindness and visual impairment; cystic fibrosis and other metabolic developmental disorders; epilepsy; fibromyalgia, chronic pain and fatigue; hearing impairment; human immunodeficiency virus and/or acquired immunodeficiency syndrome (HIV/AIDS); intellectual disabilities and related conditions; leukemias and lymphomas; migraine and chronic headache; mobility impairments; muscular dystrophy; and peripheral vascular disease. We excluded the original CCW condition category for diabetes as well as original and other chronic and potentially disabling conditions related to behavioral health or drug use conditions – specifically, alcohol use disorder; anxiety; bipolar disorder; depression; schizophrenia; opioid use disorder; and tobacco use - as they are related to current or future planned outcomes.

For the ED access measure, we ran the Elixhauser Comorbidity Software on all inpatient and outpatient ED claims included in the measure denominator. For ED visits in the ICD-9 period, we used version 3.7, and, for the ICD-10 period, we used version 2020 (AHRQ n.d.[b], n.d.[c]). From both versions of the software, we constructed the 29 individual comorbidity conditions, including acquired immune deficiency syndrome; alcohol abuse; chronic blood loss anemia; chronic pulmonary disease; congestive heart failure; coagulopathy; deficiency anemias; depression; diabetes with chronic complications; diabetes without chronic complications; drug abuse; fluid and electrolyte disorders; hypertension; hypothyroidism; liver disease; lymphoma; metastatic cancer; obesity; other neurological disorders; paralysis; peptic ulcer disease and bleeding; peripheral vascular disease; psychoses; pulmonary circulation disease; renal failure; rheumatoid arthritis/collagen vascular disease; solid tumor without metastasis; valvular disease; and weight loss. We also calculated from the software the readmissions summary score and mortality summary score. We modified both ICD-9 and ICD-10 algorithms, however, to not apply logic related to DRGs to define the condition categories because there is no DRG variable on the outpatient ED claims, and we wanted to define the condition categories consistently across the inpatient and outpatient ED claims. The DRG variable was typically used as an exclusion criterion (that is, to not flag a claim as having a specific condition). For this reason, the prevalence of condition categories might be higher in our analysis file than in other studies of hospitalized Medicare beneficiaries. Exhibit B.26 defines the rest of the covariates used in the impact regressions.

Exhibit B.26. Covariates for beneficiary- and discharge-level regression models

Covariate	Definition	Included in regressions at: Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
Demographics and enrollment characteristics						
Age category	**Calculated based on the first day observable in Medicare data for the year (that is, alive, enrolled in Parts A and B FFS Medicare, with Medicare as primary payer)**					
Age less than 65 years (omitted)[a]	= 1 if age < 65 years = 0 otherwise					
Ages 65 to 69 years (reference category)	= 1 if age >= 65 years & age <= 69 years = 0 otherwise					
Ages 70 to 74 years	= 1 if age >= 70 years & age <=74 years = 0 otherwise	X	X	X	X	X
Ages 75 to 79 years	= 1 if age >= 75 years & age <=79 years = 0 otherwise	X	X	X	X	X
Ages 80 to 84 years	= 1 if age >= 80 years & age<=84 years = 0 otherwise	X	X	X	X	X
Ages 85 years and older	= 1 if age >= 85 years = 0 otherwise	X	X	X	X	X
Sex						
Male (reference category)	= 1 if male or unknown sex = 0 if female					
Female	= 1 if female = 0 if male or unknown sex	X	X	X	X	X

		Included in regressions at:				
Covariate	Definition	Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
Race and ethnicity[b]						
White (reference category)	= 1 if RTI race variable = 1 = 0 if RTI race variable not equal to 1					
Black	= 1 if RTI race variable = 2 = 0 if RTI race variable not equal to 2	X	X	X	X	X
Hispanic	= 1 if RTI race variable = 5 = 0 if RTI race variable not equal to 5	X	X	X	X	X
Other minorities	= 1 if RTI race variable = {0,3,4,6} = 0 if RTI race variable = {1,2,5}	X	X	X	X	X
OREC and age interaction						
Not disabled or ESRD (reference category)	= 1 if OREC = disabled, ESRD, or disabled and ESRD = 0 otherwise					
Age < 65 and disabled or ESRD	= 1 if OREC = disabled, ESRD, or disabled and ESRD and age<65 years = 0 if age>65 years or age <65 and OREC = aged	X	X	X	X	X
Age >=65 and ESRD	= 1 if OREC = ESRD or disabled and ESRD and age >=65 years = 0 if OREC = aged or disabled or if age < 65 years	X	X	X	X	X
Ages 65 to 69 years and disabled	= 1 if OREC = disabled and age >= 65 and age < 70 = 0 if OREC = aged, ESRD, or disabled and ESRD or age < 65 or age >= 70	X	X	X	X	X

Covariate	Definition	Included in regressions at: Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
Ages 70 to 74 years and disabled	= 1 if OREC = disabled and age >= 70 and age < 75 = 0 if OREC = aged, ESRD, or disabled and ESRD or age < 70 or age >= 75	X	X	X	X	X
Age 75 to 79 years and disabled	= 1 if OREC = disabled and age >= 75 and age < 80 = 0 if OREC = aged, ESRD, or disabled and ESRD or age < 75 or age >= 80	X	X	X	X	X
Age 80 to 84 years and disabled	= 1 if OREC = disabled and age >= 80 and age < 85 = 0 if OREC = aged, ESRD, or disabled and ESRD or age < 80 or age >= 85	X	X	X	X	X
Age 85 years and older and disabled	= 1 if OREC = disabled and age >= 85 = 0 if OREC = aged, ESRD, or disabled and ESRD or age < 85	X	X	X	X	X
Social Vulnerability Index[c]	CDC Social Vulnerability Index merged onto individual beneficiaries at the Census tract level. Overall ranking of Census tract from 15 social factors, including poverty, lack of vehicle access, and crowded housing	X	X	X	X	X
Rural residence	= 1 if more than 50% of residents in that zip code are living in rural areas, per Census Urban and Rural classification by ZCTA = 0 if 50% or fewer residents in that zip code are living in rural areas, per Census Urban and Rural classification by ZCTA	X	X			
Education	4-category variable for whether the beneficiary reported their highest level of educational attainment as: 1) less than high school; 2) high school (reference category); 3) some college; or 4) college degree or higher		X			

Covariate	Definition	Included in regressions at: Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
CCW-related variables[d]						
CCW Condition flags	= 1 if the beneficiary had claims-based evidence of the condition in the relevant look-back period = 0 if the beneficiary had no claims-based evidence of the condition in the relevant look-back period = missing if the beneficiary was not observable in the relevant look-back period	X		X	X	X
Observable for CCW one-year look-back period[e]	= 1 if the beneficiary was observable in the one-year look-back period used for multiple CCW condition categories = 0 if the beneficiary was not observable in the one-year look-back period used for multiple CCW condition categories	X		X	X	X
Observable for CCW two-year look-back period[f]	= 1 if the beneficiary was observable in the two-year look-back period used for multiple CCW condition categories = 0 if the beneficiary was not observable in the two-year look-back period used for multiple CCW condition categories	X		X	X	X
Observable for CCW three-year look-back period[g]	= 1 if the beneficiary was observable in the three-year look-back period used for the CCW Alzheimer's and related disorders and senile dementia condition category = 0 if the beneficiary was not observable in the three-year look-back period used for the CCW Alzheimer's and related disorders and senile dementia condition category	X		X	X	X

Covariate	Definition	Included in regressions at: Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
New enrollee flag	= 1 if the beneficiary was not observable in all months of the prior year = 0 if the beneficiary was observable in all months of the prior year	X		X	X	X
Three or more physical health conditions	=1 if the number of individual CCW condition categories related to physical health is greater than or equal to 3, excluding hypertension and hyperlipidemia =0 if the number of individual HCC condition categories related to physical health is 0, 1, or 2, excluding hypertension and hyperlipidemia = missing if the beneficiary was not observable in all months of the prior year	X		X	X	X
Elixhauser-like variables[d, h]						
Condition categories	= 1 if the beneficiary had evidence of the condition on the ED visit claim = 0 if the beneficiary had no evidence of the condition on the ED visit claim					X
Mortality summary score	A score constructed by the software that weights each condition based on how it contributes to predicting mortality and then sums across all categories to get a single score					X
Readmission summary score	A score constructed by the software that weights each condition based on how it contributes to predicting readmissions and then sums across all categories to get a single score					X

		Included in regressions at:				
Covariate	**Definition**	**Beneficiary level: Claims measures**	**Beneficiary level: CAHPS measures**	**Discharge level: 30-day unplanned readmission**	**Discharge level: Timely follow-up**	**ED visit level: ED access**
Other measure of health status						
Self-reported health status	3-category variable for whether the beneficiary self-reported their health status as 1) poor/fair; 2) good (reference category) or 3) very good or excellent		X			
Episode of care-related variables						
Chronic condition category for acute exacerbation follow-up[i]	= Asthma or hypertension or coronary artery disease or heart failure or chronic obstructive pulmonary disease or diabetes if the principal diagnosis code is sufficient to diagnose the condition or if the principal diagnosis code is related to the condition and a secondary diagnosis on the claim is sufficient, per IMPAQ technical specifications				X	
Index admission category for 30-day unplanned readmission[j]	=Surgical, cardiorespiratory, cardiovascular, neurology, or medicine based on the procedure codes and principal diagnosis, per CMS/Yale technical specifications			X		
COVID-19 variables[j,k]						
COVID-19 inpatient stay	Two variables, one measured for 2020 and another separately for 2021, but both defined as: =1 if beneficiary had at least one inpatient admission with a diagnosis of COVID-19 =0 if beneficiary had no inpatient admissions with a diagnosis of COVID-19	X		X	X	

Covariate	Definition	Included in regressions at: Beneficiary level: Claims measures	Beneficiary level: CAHPS measures	Discharge level: 30-day unplanned readmission	Discharge level: Timely follow-up	ED visit level: ED access
COVID-19 ED or observation visit	Two variables, one measured for 2020 and another separately for 2021, but both defined as: =1 if beneficiary had at least one ED visit or observation stay with a diagnosis of COVID-19 =0 if beneficiary had no ED visits or observation stays with a diagnosis of COVID-19	X		X	X	

[a] Age less than 65 years was collinear with Age < 65 and OREC = disabled or ESRD.

[b] We combined other minorities into a single category for regression due to the small number of beneficiaries who meet this definition in Maryland.

[c] We allowed the SVI summary score to change over time for each census tract. Specifically, for calendar years 2011 through 2013, we used the 2012 SVI summary score. For calendar years 2014 and 2015, we used the 2014 SVI summary score. For calendar years 2016 and 2017, we used the 2016 SVI summary score. For calendar years 2018 through 2022, we used the 2018 SVI summary score. We did not use the 2020 SVI summary score because the measure definition changed from the prior years' SVI.

[d] See text in Section B.2.2 for a list of conditions included.

[e] The following original CCW condition categories have a one-year look-back period: acquired hypothyroidism; acute myocardial infarction; anemia; asthma; atrial fibrillation; benign prostatic hyperplasia; cancer – breast; cancer – colorectal; cancer – endometrial; cancer – lung; cancer – prostate; cataract; chronic obstructive pulmonary disease and bronchiectasis; glaucoma; hip/pelvic fracture; hyperlipidemia; hypertension; osteoporosis; and stroke/transient ischemic attack.

[f] The following original CCW condition categories have a two-year look-back period: chronic kidney disease; heart failure; ischemic heart disease; rheumatoid arthritis/osteoarthritis. In addition, all of the other chronic and potentially disabling conditions have a two-year look-back period, including: blindness and visual impairment; cystic fibrosis and other metabolic developmental disorders; epilepsy; fibromyalgia, chronic pain and fatigue; hearing impairment; human immunodeficiency virus and/or acquired immunodeficiency syndrome (HIV/AIDS); intellectual disabilities and related conditions; leukemias and lymphomas; migraine and chronic headache; mobility impairments; muscular dystrophy; and peripheral vascular disease.

[g] The CCW condition category for Alzheimer's disease and related disorders or senile dementia has a three-year look-back period.

[h] We refer to these as "Elixhauser-like" because we did not apply DRG-related exclusion criteria to relevant condition categories. Thus, the prevalence of conditions might be higher in our population than other published analyses of Medicare FFS populations.

[i] See the IMPAQ Health (2018) specifications for details on the chronic condition category assignment

j See the specifications for 30-day unplanned readmission developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2020) for details on index admission category assignment

[k] The COVID-19 variables were included in sensitivity analyses and not the main results. See Appendix C.

CCW = Chronic Conditions Warehouse; CAHPS = Consumer Assessment of Healthcare Providers and Systems; ED = emergency department; HCC = Hierarchical Condition Category; MSBF = Medicare Beneficiary Summary File; OREC = Original Reason for Entitlement Code; RTI = Research Triangle Institute; TIA = transient ischemic attack; VRDC = Virtual Research Data Center; ZCTA = Zip Code Tabulation Area.

B.2.2c. Hospital-year level files for impact analyses of patients' rating of their hospital

We fit regression models at the hospital-year level for the HCAHPS-based patient experience measure to assess impacts of the Maryland Model on patients' overall ratings of their hospital stays (see Appendix B.1.6. for more details on regression specifications). The hospital-year-level analytic file is constructed from publicly available data obtained from CMS's Hospital Compare website; it contains one observation per hospital per year (2011–2020) for all Medicare-certified hospitals, identified by their CMS Certification number (CMS 2021b). To control for differences in case-mix over time between Maryland and comparison group hospitals, we also merged on hospital case mix index from CMS's Inpatient Prospective Payment System public use files in each year by CMS Certification number (CMS n.d.; CMS 2022c). Hospital case mix index is calculated from the Medicare Severity Diagnosis-Related Groups weight for each stay, which reflects the mean severity of all stays at the hospital during the year.

To ensure we had an appropriate analytic sample on which to estimate impacts, we made several restrictions to the full list of hospitals in Hospital Compare for our primary regression models, including the following:

- We dropped hospitals without CAHPS scores in a given year. Hospitals might appear in the data without a score because they are exempted (that is, not subject to the Inpatient Prospective Payment System) or because they failed to reach the minimum number of survey responses to avoid suppression in the data.
- We dropped hospitals not located in a PUMA that is part of our comparison group (that is, received zero matching weight) or that did not have any address information (street or ZIP code). Though the data are at the hospital-year level, we continue to weight hospitals in the rest of the nation using our PUMA-level matching weight to ensure we are using the same comparison group as we are for other outcomes. This means we had to assign hospitals to PUMAs using address information available in Hospital Compare. We used GIS mapping software to identify addresses and place hospitals into PUMAs.
- We dropped hospitals that did not have any Medicare FFS discharges in 2013 or, if not in our data in 2013, the year after they first appear in Hospital Compare. Our final regression weight for patient hospital ratings multiplies the PUMA matching weight by a weight representing the size of the hospital, measured by total number of discharges in 2013. Hospitals that do not have any Medicare FFS discharges in claims could not be assigned a weight. We fixed the hospital size weight based on hospital discharges in 2013 (the year before the MDAPM period began) because the Maryland Model might impact the number of discharges in years after the intervention began. For hospitals not in our data in 2013 (for example, new hospitals in later years), we used the number of discharges in the year after they first appear as their size weight in all years.

- We dropped hospitals with missing information for case mix index in all years that appear in our analysis.[29] In our primary models, we want to control for patient case mix because it is possible that the Maryland Model and its incentives to move care out of the hospital could change the case mix in Maryland hospitals relative to hospitals in the comparison group. A consequence of this restriction is that all Critical Access Hospitals (CAHs) drop out of our primary sample because they are not subject to the Inpatient Prospective Payment System and therefore do not have case mix index information in any year. There are no official CAHs in Maryland, but several hospitals, particularly in rural areas, function similarly to how CAHs do in the rest of the nation. We tested models that retained CAHs in our regressions by dropping case mix index and found our results to be very consistent with our main findings.
- We dropped hospitals that only appeared in a single year in our data (after making the restrictions above). Our primary models chose not to include hospital-level fixed effects in an effort to avoid over-controlling for potential effects of the Maryland Model on hospital closure. Hospitals that only appear once are likely data anomalies (for example, a hospital converting to a different type or newly merging with another hospital). We don't believe these single data-point hospitals meaningfully contribute to our analysis, and they could introduce bias if mismeasured differentially in Maryland and the comparison group.
- Finally, in Maryland, we restrict observations in our analysis to the list of hospitals that are part of the 2013 or 2019 State Agreements (HSCRC 2018). For this report, after applying the criteria above, this did not remove any additional hospitals from our analytic sample but may come into play in future years.

The final hospital analytic panel contains 761 unique hospitals over 11 years that meet the above criteria.

B.3. Supplemental methods and findings for hospital responses to model incentives and supports

B.3.1. Hospital site visits methods and key findings

B.3.1a. Methods

- We conducted virtual site visits with seven general acute care hospitals. For each site visit, we conducted semi-structured interviews with seven to 10 respondents using a video conferencing platform individually or in small groups.
- We aimed for the hospitals we selected to represent all Maryland acute care hospitals in terms of (1) program participation in Regional Partnership Catalyst Grants, the Episode Care Improvement Program (ECIP), and Care Transformation Initiatives (CTIs); (2) whether the

[29] A small number of hospitals had case-mix index information for some years but not others. We chose to impute the missing case mix index in these cases using the mean for that hospital in the years we had valid observations. We then included a missing indicator flag for missing case mix index in our regressions.

hospital is a rural hospital that participated in the Total Patient Revenue model[30] or is independent from a health system; (3) region where the hospital is located; and (4) social vulnerability of the hospital's patient population, as defined by SVI score (CDC n.d.). Specifically, we divided hospitals into terciles based on their SVI, which is measure of an area's vulnerability to disaster (including natural disasters and infectious disease outbreaks) and includes 15 social factors, including poverty, lack of vehicle access, and crowded housing. We included at least two hospitals from each tercile.[31]

- The first interview we conducted at each hospital included the hospital's chief executive officer/president and chief financial officer. We then used a snowball sampling approach to identify additional respondents and conducted subsequent interviews with a range of staff involved in and knowledgeable about the hospital's efforts to respond to model incentives during the MD TCOC period. All interviews were recorded and professionally transcribed.
- We contemporaneously summarized each interview using a structured note-taking template. To identify key themes, two team members independently reviewed the note-taking template and all interview transcripts for each hospital and consolidated these materials into a single summary of key findings for the hospital. We then looked across the hospital summaries to identify common experiences, facilitators, and challenges as well as notable differences in hospital experiences.

B.3.1b. Key findings

Overall model incentives

- All-payer global budgets and the quality adjustments to them serve as the primary model incentive during the MD TCOC period that drove changes in hospital care delivery. One hospital described global budgets as a "great motivator to reduce unnecessary admissions" and keep patients out of the hospital.
- No hospitals mentioned the Medicare Performance Adjustment as a major incentive for driving changes in hospital care delivery. Leaders from two different hospitals noted that the Medicare Performance Adjustment helped prompt conversations among leaders about total cost of care, not just hospital costs. In one case, this broader focus had some influence on the hospital's care delivery, especially focusing care management services on the hospital's patients with highest cost of care from all settings, not just the hospital.
- Hospitals made care changes both because the incentives under the model supported them and because the changes were consistent with the hospitals' missions.

[30] The Total Patient Revenue model was implemented in 2010 at eight rural hospitals in Maryland. Similar to the Maryland model, participating rural hospitals were given fixed global operating budgets to provide care for the people in their service area. For additional information about the Total Patient Revenue model, see Pines et al. (2019).

[31] Specifically, for each hospital, we (1) identified the Medicare FFS beneficiaries discharged from the hospital in 2021, (2) identified the SVI score for the census tract where the beneficiary lives (based on residence as recorded in the Medicare Enrollment Database), and (3) found the average SVI score across all beneficiaries discharged from the hospital.

- Although all hospitals we interviewed agreed with the conceptual direction of the model during the MD TCOC period, most hospitals perceived challenges to some operational aspects of the model. For example, respondents from several hospitals felt that the model is unnecessarily complex and that the methodology of specific incentives and supports change too frequently, making it hard to develop a culture around quality improvement, to plan strategically, and to develop habits that can lead to specific improvements.
- One system described how HSCRC provided their health system with a financial incentive to convert hospitals to free-standing emergency departments when the demand for inpatient services had dissipated. When converting a hospital to a free-standing emergency department, the health system can keep some of the global budget associated with the historic services.[32] The system described these conversions as good for total cost of care because they reduce the size of the global budgets. The system also described the conversions as good for population health, in part because the system agrees with HSCRC to make certain investments to serve the community as a condition for retaining some of the former global budget. Converted facilities can vary in the services they offer but often provide emergency care, diagnostic services, and observation stays. Clinics that were previously offered at the hospital are often converted to non-hospital-based clinic meaning their services are then paid for outside of hospital global budgets.

Quality-adjusted hospital global budgets

- Hospitals implemented a range of strategies to be successful under global budgets (Exhibit B.27). Many of these programs began during the MDAPM period (2014 to 2018) but have continued to evolve over time. Two hospitals in our sample noted that the system had already "wrung out a lot of potentially avoidable utilization from hospitals" before the MD TCOC period and questioned how much additional acute care could be avoided.

[32] When a hospital converts to a free-standing medical facility, HSCRC continues to set an all-payer global budget for the facility.

Exhibit B.27. Hospitals' strategies across the care continuum in response to quality-adjusted global budgets

Strategy	Examples from site visits
Reducing preventable acute care	
Improving hospital care delivery	• Enhanced discharge planning • Multidisciplinary rounding to align providers on discharge plans • Schedule follow-up appointments before discharge
Coordinating care with non-hospital providers	• Partner with post-acute care providers to improve their quality and offer medical directors in certain cases to strengthen collaboration • Partner with primary care to promote advance directive, establish care plans, and encourage preventative care • Provide care management to high-risk people after hospital discharge to, for example, ensure they have follow-up appointments and access to medications
Addressing social determinants of health	• Invest in community health workers to enhance outreach, education, and support for underserved populations • Implement Diabetes Prevention Program to promote healthy lifestyle changes • Screen for various health conditions in non-traditional settings, such as grocery stores and mobile health clinics • Partner with community organizations to address health-related social needs, such as housing with wraparound services
Shifting care to lower-acuity settings	
Investments in alternatives to the ED	• Open urgent care centers • Partner with or open behavioral health crisis centers • Partner with Federally Qualified Health Centers for dental care • Partner with emergency medical services to visit the homes of frequent utilizers of the 911 system and identify solutions
Investments in alternatives to admissions	• ED care management to find alternatives to hospital admissions for low-acuity concerns, such as connecting people with community resources to address social issues • Transition appropriate cases from the hospital to ambulatory surgical centers or other non-hospital care settings, such as urgent care centers
Limiting medical overuse, not providing low-value services, or both	
Hospice	• Expand hospice and palliative care teams to minimize acute care utilization at end of life • Establish joint ventures with hospice providers to offer hospice beds within the hospital setting (but outside of global budgets)

Source: Site visits with hospitals.

ED = emergency department.

- The quality adjustments to global budgets have prompted care changes primarily aligned with federal initiatives, although a few hospitals mentioned that some adjustments were unique or larger in Maryland. One hospital described seeking win-win situations across quality programs and other model incentives. For example, this hospital concentrated on reducing hospital-acquired infections, which it believed contributed to improved performance on metrics such as Hospital-Acquired Conditions, Quality-Based Reimbursement, readmissions, and the global budget incentive. However, several respondents from various hospitals acknowledged that the complexity and evolving nature of the quality adjustments pose a challenge and could diminish their impact (as mentioned earlier).

- All tertiary care hospitals we interviewed felt global budgets are ill-suited to providing high acuity and specialized care.
  - The main challenge lies in the fact that hospitals are rewarded under global budgets for reducing avoidable utilization. These hospitals argue, however, that most of their volume cannot or should not be avoided. Although they also provide some community care, they contend that even when they reduce avoidable utilization in those cases, it does not result in retained revenue because of "a relatively unlimited demand signal." (That is, efforts to reduce preventable hospital use do not actually lower hospital use because prevented stays are replaced by others waiting to use the hospital.) These hospitals all reported turning away patients in need of their services daily.
  - Furthermore, these hospitals reported that community hospitals responded to global budgets by strategically offloading volume, which was then absorbed by tertiary care centers. For example, one hospital explained that many community hospitals cannot afford to have neurologists on call and therefore end up sending all neurology cases to their hospital. Although some of the volume they received might benefit patients in terms of quality, the hospitals argued that a significant portion of that volume would be better served in the community. This influx of volume not only strains their beds and resources but also negatively affects patients who now must travel longer distances.
  - The innovation adjustment and market shift adjustment were considered insufficient to overcome these challenges. Instead, these hospitals suggested that their tertiary care be excluded from global budgets and paid on an FFS basis instead.
- Several hospitals reported that global budgets present challenges in terms of offering new services or specialty care programs. Hospitals explained that they are not reimbursed for the full cost of implementing new specialty care and that they need to have a certain volume of patients receiving that care to maintain quality. Regarding the latter point, interviews from one hospital noted that out-of-state Medicaid does not want to pay Maryland rates, which limits the population of patients the hospital can serve.
- Hospitals reported that these challenges, in turn, impact the ability to attract and retain academic physicians and to provide access to novel technologies and services to meet community need. As one respondent shared, "On the one hand I know the intent is to control unnecessary growth, which is needed to control expenses. But on the other hand, I think it holds Maryland back a little bit in appropriately growing services and service lines, matched up with community need."
- Hospitals reported that global budgets provided financial stability and predictability through COVID-19, especially at the onset of the pandemic. Furthermore, a few hospitals noted that global budgets facilitated the acceleration of certain types of care delivery that helped meet patient needs during the pandemic. Examples include shifting cases to outpatient settings, implementing new programs such as patient remote monitoring, and forging stronger partnerships with post-acute care providers. Hospitals also shared that COVID disrupted care

delivery changes aimed at reducing avoidable hospital use because of staff shortages or redeployments.

Non-hospital spending

- Hospitals reported that model incentives to curb non-hospital spending during the MD TCOC period were too small to change care. In addition, most hospitals noted challenges in trying to curb non-hospital spending, including the fact that most non-hospital providers are paid FFS and therefore incentivized to increase volume.
- Many hospitals reported engaging in collaborations with non-hospital providers, although these efforts were primarily focused on reducing hospital utilization rather than specifically limiting growth in non-hospital spending. For instance, hospitals are forming partnerships with primary care providers to reduce readmissions or ED visits.
- Hospitals' initiatives to decrease non-hospital spending predominantly rely on clinically integrated networks that were established before the MD TCOC period and designed in part to support value-based contracts (often the Medicare Shared Savings Program). Hospitals, or their systems, acting as Care Transformation Organizations during the TCOC period, often provide Care Transformation Organization supports and services through their clinically integrated networks.
- Hospitals noted that it was harder to control non-hospital spending in competitive markets where patients have the freedom to seek care anywhere. Conversely, it is easier for hospitals that serve as sole community providers or have ownership across different parts of the care continuum.
- Hospital noted specific challenges with the Medicare Performance Adjustment as a mechanism for controlling total cost of care. Several hospitals highlighted the challenge of being responsible for all patients in their service area, regardless of whether those patients seek care at local hospital in a given year. Hospitals are limited in their ability to influence care patterns and costs for patients who do not use the hospital. One tertiary care center emphasized that it serves patients from all over the state and then returns them to their respective zip codes. It preferred the concept of managing the total cost of an episode of care.
- To align incentives with non-hospital providers and promote true accountability for total cost of care, one rural hospital proposed a global budget 2.0 that encompasses both hospital and non-hospital spending.

ECIP

- Hospitals chose to participate in ECIP to help providers obtain advanced alternative payment model status[33] and improve coordination with post-acute care partners. None of the hospitals

[33] Participants in advanced alternative payment models that are part of CMS's Quality Payment Program could earn a 5% increase in their covered professional services payments through 2022. See https://qpp.cms.gov/apms/overview.

mentioned the incentive payments that hospitals could earn under the program as a motivating factor for participation.

- Several hospitals found ECIP helpful in improving coordination with post-acute care partners. For example, one ECIP care partner (a SNF) mentioned that it received a stoplight tool from the hospital to identify early symptoms of declining health, provide treatment on site at the SNF, and avoid readmissions. It was not always clear whether these collaborations would have occurred without ECIP, as some collaborations appeared to also be motivated by COVID-19 and other hospital or system priorities.
- Some hospitals found the ECIP program to be administratively burdensome. Several hospitals planned to discontinue their participation, with some of them repacking their work into a CTI because it had fewer administrative components. A few hospitals expressed an intention to expand their participation in the ECIP program.

CTIs

- CTIs had broad participation, with all hospitals in our sample participating in one or more CTIs. All the hospitals we interviewed reported packaging existing initiatives as CTIs. One respondent stated that their hospital "selected things that either we were just getting ready to implement anyway at the time the CTI program [was implemented] ...or things that we had really just started doing in the last couple of years, but we [thought] would show improvement in those utilization metrics."
- Several hospitals noted that although the CTI incentive structure encouraged participation, it also discouraged hospitals from investing additional resources. Hospitals understood that they needed to actively engage to have a chance of receiving an incentive payment. But even if a hospital succeeded under the program, it could still end up owing money because performance was relative to other hospital participants, so it was not reasonable to invest new resources. One hospital described CTIs as an "unknowable and unwinnable game."

Regional Partnership Catalyst Grants

- Respondents credited the Regional Partnership Catalyst Grants for fostering an "unprecedented level of collaboration" among hospitals in the state, attributing this to the allocation of real dollars to address important issues.
- Hospitals are using diabetes partnership funds to improve referral and enrollment in the Diabetes Prevention Program and the Diabetes Self-Management Training program. This includes community marketing, establishing referral networks, facilitating prompt class initiation after referrals, and providing support services to overcome social barriers such as transportation. Hospitals anticipate shifting focus to completion rates and sustainability in the future.
- Hospitals involved in diabetes partnerships highlighted challenges with retention and billing. Hospitals faced obstacles with the time commitment and in-person attendance requirements for classes. Some grants explored remote options, but they cannot be billed for, potentially

affecting long-term sustainability. Notably, one partnership we interviewed had its grant funding terminated because of low enrollment in courses.

- Hospitals participating in behavioral health grants are in the process of establishing call centers, mobile crisis teams, and crisis stabilization centers, aiming to divert behavioral health patients from the ED and hospital care. The crisis stabilization centers were in their early stages or not yet operating at full capacity during the interviews, but there was significant enthusiasm for their role in addressing a critical need in the state.

Health equity

- We did not explicitly ask hospital respondents about the Statewide Integrated Health Improvement Strategy (SIHIS), and no respondents mentioned the SIHIS when discussing their motivation for implementing changes to improve health equity.
- Hospitals shared various examples of their approaches to addressing disparities in health outcomes, often highlighting an increased focus on analyzing data by subgroups and implementing improvement projects with specific change tactics identified to address the disparity. For example, one hospital that is part of a health system mentioned that the health system analyzes available data to examine differences in health outcomes based on variables such as race, gender, neighborhood, and payer. When the system identifies a disparity, it works with a hospital team to develop quality improvement initiatives to address those disparities. It was not apparent, however, that any of these efforts were directly linked to the model incentives or supports.
- Hospitals also emphasized their initiatives to address health-related social needs by focusing on non-medical causes of disparities in health outcomes, which have the potential to reduce health disparities over time and align with the core incentive in the model to reduce avoidable acute care. For instance, hospitals highlighted strategies such as improving transportation to appointments and reaching out to isolated patients to decrease ED utilization. Another hospital mentioned its provision of supportive housing with wraparound services, noting that "the Maryland Model does give us incentives to do these kinds of collective, big investments in social determinants of health issues, that don't exist elsewhere."

B.3.2. Hospital survey methods and key findings

B.3.2a. Methods

We fielded a web survey (Appendix B.3.2.d) to hospitals using QuestionPro from July 28, 2022, to October 13, 2022. The survey consisted of five domains: Incentives and opportunities in the MD TCOC Model (three questions), Hospital global budgets (14 questions), Non-hospital spending (seven questions), Participation in ECIP and CTIs (five questions), and Quality of care (eight questions). We fielded the survey to 44 hospitals, which includes all hospitals under global budgets that provide acute care. We sent the survey to hospital chief financial officers, encouraging them to reach out to others at the hospital for further input. We did not survey the seven free-standing emergency rooms or facilities that are subject to all-payer global budgets but do not have the potential to participate in all model components (for example, they cannot participate in ECIP because it requires an inpatient admission). Although some content from the hospital survey would be directly relevant to these facility types, it was not in our budget to create a unique survey for these facilities. The survey achieved a 100% response rate.

B.3.2b. Key findings

We've organized the key findings by the order they appear in the survey. Tables with results for all survey questions are in section B.3.2.c.

Incentives and opportunities in the MD TCOC Model

- All-payer global budgets and the quality adjustments to them have the greatest influence on investments that hospitals have made since the start of the MD TCOC period to change care delivery, with 74% of hospitals reporting that these incentives influenced their investments a lot (A1). Investments refer to contributions of staff, infrastructure, or other resources to support care delivery changes. Respondents chose what "a lot" meant in their context, choosing between four options for how much the incentive influenced their care delivery investment: "Not at all," "a little," "some," or "a lot."
- In contrast, newer episode-based model incentives had less influence on hospitals' investments, with more than half of hospitals (58%) reporting CTIs have little or no influence on their investments, and 79% of hospitals reporting ECIP had little or no influence on their investments (A1).
- About half of hospitals reported that the model incentives are clear, allowing hospitals to make informed decisions about how to respond to them (A2). For hospitals that are part of a system, 87% reported that these decisions were made either fully or partly by the system (A3).

Hospital global budgets

- To do well financially under global budgets, many hospitals reported investing a lot in improving performance on quality measures and reducing preventable hospital use (84% and 82%, respectively). Around half of hospitals also reported investing a lot into shifting care to lower-acuity settings when appropriate (B1).
- To reduce preventable hospital use, many hospitals reported investing a lot in (1) improving hospital care delivery (80%) and (2) coordinating medical or behavioral health care with non-hospital providers (70%). Fewer hospitals reported investing a lot into addressing social determinants of health through hospital staff (32%) or by coordinating with non-hospital providers or community-based organizations (30%) (B4).
- To shift care to lower-acuity settings, many hospitals reported investing a lot in (1) shifted ED care to non-hospital settings (52%) and (2) shifted hospitalized patients to SNFs or other post-acute care more quickly (52%). Around one-third of hospitals were also investing a lot

in shifting care within their hospital to lower- versus higher-intensity services (30%), and 12% were shifting surgeries or other planned medical care to non-hospital settings (B6).

- In the summer of 2022 when we fielded the survey, hospitals reported that (1) high inflation and (2) the COVID-19 pandemic (86 and 81%, respectively) were a great deal of a barrier in their ongoing efforts to succeed under global budgets. In addition, 55% of hospitals noted that difference in payment and incentives between hospitals and other providers were a great deal of a barrier (B14).
- Hospitals had mixed perspectives on their ability to maintain financial viability under the model (B13). Around half of hospitals agreed or strongly agreed they could operate in a financially viable way under the model, but 39% of hospitals disagreed or strongly disagreed with this sentiment. The timing of the survey likely influenced this finding, as it coincided with a period marked by high inflation and the ongoing impact of the COVID-19 pandemic, which were widely viewed by hospitals as dominant factors affecting their operations under a global budget. For instance, one hospital stated in the free text response, "COVID and the post COVID inflation has put a significant burden on the financial health of the hospital. Inflation amendments to our rates tends to lag actual inflation so we are experiencing financial pressures when hospital rates have not yet increased for this unusual spike in inflation (especially wage inflation in post COVID labor market)."

Non-hospital spending

- Although many hospitals (57%) reported feeling they have some influence on non-hospital spending among Medicare beneficiaries in their area, no hospital reported having a lot of influence in this area. Furthermore, 43% reported feeling they have little or no influence on non-hospital spending (C1).
- Hospitals' investments to limit growth in non-hospital Medicare spending varied; around one-third of hospitals (30%) reported investing a lot, another one-third (36%) reported investing some, and the final one-third (34%) reported investing little to none (C2). However, 60%of hospitals reported their investment increased since the model began (C3).
- Among various MD TCOC Model and non-model incentives and supports, hospitals most commonly reported payments for being part of a CTO (31%) and contracts with Accountable Care Organizations (17%) as highly encouraging factors to limit the growth of non-hospital spending. In comparison, Regional Partnership Catalyst Grants and the Medicare Performance Adjustment were less influential, with fewer hospitals reporting these incentives encouraged them a lot to limit growth in non-hospital spending. CTIs and ECIP were deemed the least influential, with only about 2% of hospitals reporting these incentives encouraged them a lot (C4).
- To curb the growth of non-hospital spending, hospitals reported the biggest investments in improving quality of hospital care to prevent post-discharge complications (81%) and partnering with primary care or other providers to improve management of chronic conditions (65%). These areas might also help them do well under global budgets. Although

less common, some hospitals were still investing a lot in other strategies to curb the growth in non-hospital spending, including the following (C5):

- Discharging patients to lower-cost-of-care settings (for example, to home health instead of to SNFs) (44%)
- Partnering with community-based organizations to address social determinants of health and prevent chronic conditions (38%)
- Partnering with primary care or other providers to increase adherence to clinical care guidelines (which could limit provision of low-value services) (28%)
- Partnering with primary care providers, specialty providers, or both to reduce duplication of imaging or testing (9%)

- Hospitals reported that (1) challenges operating during COVID-19 and (2) the inability to affect patients you rarely or never interact with were a great deal of a barrier in their efforts to limit non-hospital spending for Medicare beneficiaries in their hospital's service area (88 and 67%, respectively). In addition, 47% of hospitals noted that conflicting incentives (from global budgets) for hospitals to shift care outside of the hospital, which increases non-hospital spending, were a great deal of a barrier (C7).

Participation in ECIP and CTIs

- Among hospitals participating in ECIP, more than half reported that limited downside risk and the ability of care partners to qualify for CMS bonus payments for participating in Advanced Alternative Payment Models were very important factors in their participation (73 and 64%, respectively). Yet only 5% of hospitals indicated that the ability to share incentives with care partners was very important, suggesting that the waiver of anti-kickback statutes, which is part of the program, holds minimal importance for hospitals (D1).
- Among hospitals participating in CTIs, around half reported that the interventions delivered under CTIs should also help them perform well under global budgets, the opportunity for the hospital to earn incentive payments, and the chance to address an area of high service use and spending were very important factors in their participation (D4). The finding that many hospitals perceive CTIs as beneficial for succeeding under global budgets might reflect that CTIs are based on standardized hospital spending and therefore hospitals can succeed in these programs by reducing preventable hospital use.

Quality of care

- Around two-thirds of hospitals reported engaging a lot in efforts to improve equity in health outcomes (E6), but hospitals were split on whether these efforts are encouraged by the model, with 14% saying it was encouraged "a lot" by the model and around one-quarter saying "not at all." (E7).

B.3.2c. Tables of results for all survey questions

Section A. Incentives and opportunities in the MD TCOC Model

	Percent of hospitals
A1_a. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: All-payer hospital global budgets	
A lot	74.4
Some	7.0
A little	18.6
Not at all	0.0
N	43
A1_b. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: Quality adjustments to global budgets	
A lot	74.4
Some	20.9
A little	4.7
Not at all	0.0
N	43
A1_c. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: The Medicare Performance Adjustment	
A lot	39.5
Some	20.9
A little	37.2
Not at all	2.3
N	43
A1_d. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: Incentives for efficient episodes of care under ECIP	
A lot	2.3
Some	18.6
A little	53.5
Not at all	25.6
N	43

	Percent of hospitals
A1_e. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: Incentives for efficient episodes of care under CTIs	
A lot	11.6
Some	30.2
A little	51.2
Not at all	7.0
N	43
A1_f. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: Regional Partnership Grants	
A lot	23.3
Some	25.6
A little	44.2
Not at all	7.0
N	43
A1_g. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model? Investment refers to contributions of staff, infrastructure, or other resources by your hospital: Payments for being part of a Care Transformation Organization under MDPCP	
A lot	46.5
Some	30.2
A little	11.6
Not at all	11.6
N	43
A2. Please indicate how much you agree or disagree with the following statement. The incentives under the MD TCOC Model are clear, allowing our hospital to make informed decisions about how to respond to them.	
Strongly agree	18.6
Agree	34.9
Neither agree nor disagree	11.6
Disagree	30.2
Strongly disagree	4.7
N	43
A3. [For hospitals that are part of a health system] To what extent are decisions about how your hospital responds to MD TCOC Model incentives made by your hospital's health system?	
Fully - Decisions are centralized within the health system	42.1
Partly - System and hospital leadership jointly decide	44.7
Little - Leaders at our hospital largely make decisions on our own, with some input from the health system	10.5
None - Our hospital makes decisions independently, with no input from the health system	2.6
N	38

Section B. Hospital global budgets

	Percent of hospitals
B1_a. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Reducing preventable hospital use (that is, hospital use that is unplanned and may be prevented through improved care, care coordination, or effective community-based care)	
A lot	81.8
Some	18.2
A little	0.0
None	0.0
N	44
B1_b. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Shifting care to lower-acuity settings (for example, ambulatory care settings) when appropriate	
A lot	47.7
Some	47.7
A little	4.5
None	0.0
N	44
B1_c. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Limiting medical overuse, not providing low-value services, or both	
A lot	36.4
Some	47.7
A little	11.4
None	4.5
N	44
B1_d. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Increasing your hospital's global budget via the Market Shift Adjustment	
A lot	15.9
Some	25.0
A little	45.5
None	13.6
N	44
B1_e. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Improving performance on quality measures that affect hospital budgets	
A lot	84.1
Some	13.6
A little	2.3
None	0.0
N	44

	Percent of hospitals
B1_f. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget? Other (please specify)	
A lot	20.0
Some	2.9
A little	2.9
None	74.3
N	35
B2_a. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Reducing preventable hospital use	
Investment has increased	55.8
Investment has stayed the same	44.2
Investment has decreased	0.0
N	43
B2_b. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Shifting care to lower-acuity settings when appropriate	
Investment has increased	67.4
Investment has stayed the same	32.6
Investment has decreased	0.0
N	43
B2_c. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Limiting medical overuse, not providing low-value services, or both	
Investment has increased	41.9
Investment has stayed the same	58.1
Investment has decreased	0.0
N	43
B2_d. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Increasing your hospital's global budget via the Market Shift Adjustment	
Investment has increased	19.0
Investment has stayed the same	81.0
Investment has decreased	0.0
N	42
B2_e. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Improving performance on quality measures that affect hospital budgets	
Investment has increased	74.4
Investment has stayed the same	25.6
Investment has decreased	0.0
N	43

	Percent of hospitals
B2_f. How much has your hospital's investment in each of the following strategies changed since the 2019 implementation of the MD TCOC Model? Other (please specify)	
Investment has increased	88.9
Investment has stayed the same	11.1
Investment has decreased	0.0
N	9
B3. What type(s) of hospital use is your hospital primarily trying to reduce? (select all that apply)	
Emergency department care	84.1
Inpatient care	95.5
Outpatient care (not emergency department)	27.3
N	44
B4_a. How much of an investment is your hospital currently making in each of the following approaches to reduce preventable hospital use? Improving hospital care delivery (for example, discharge planning)	
A lot	79.5
Some	20.5
A little	0.0
None	0.0
N	44
B4_b. How much of an investment is your hospital currently making in each of the following approaches to reduce preventable hospital use? Coordinating medical or behavioral health care with non-hospital providers (for example, primary care, post-acute care)	
A lot	70.5
Some	18.2
A little	11.4
None	0.0
N	44
B4_c. How much of an investment is your hospital currently making in each of the following approaches to reduce preventable hospital use? Addressing social determinants of health using hospital staff	
A lot	31.8
Some	54.5
A little	13.6
None	0.0
N	44
B4_d. How much of an investment is your hospital currently making in each of the following approaches to reduce preventable hospital use? Addressing social determinants of health by coordinating with non-hospital providers or community-based organizations	
A lot	29.5
Some	56.8
A little	4.5
None	9.1
N	44

	Percent of hospitals
B4_e. How much of an investment is your hospital currently making in each of the following approaches to reduce preventable hospital use? Other (please specify)	
A lot	17.2
Some	3.4
A little	0.0
None	79.3
N	29
B6_a. How much of an investment is your hospital currently making in each of the following approaches to shift care to lower-acuity setting when appropriate? Shifting emergency department care to non-hospital settings (for example, primary care, urgent care)	
A lot	52.4
Some	33.3
A little	7.1
None	7.1
N	42
B6_b. How much of an investment is your hospital currently making in each of the following approaches to shift care to lower-acuity setting when appropriate? Shifting surgeries or other planned medical care to non-hospital settings (for example, ambulatory surgical centers, dialysis clinics)	
A lot	11.9
Some	45.2
A little	35.7
None	7.1
N	42
B6_c. How much of an investment is your hospital currently making in each of the following approaches to shift care to lower-acuity setting when appropriate? Shifting hospitalized patients to skilled nursing facilities or other post-acute care more quickly (that is, reducing length of stay), when appropriate	
A lot	52.4
Some	19.0
A little	23.8
None	4.8
N	42
B6_d. How much of an investment is your hospital currently making in each of the following approaches to shift care to lower-acuity setting when appropriate? Shifting care within your hospital to higher versus lower intensity services, when appropriate (for example, out of the intensive care unit)	
A lot	30.0
Some	27.5
A little	35.0
None	7.5
N	40

	Percent of hospitals
B6_e. How much of an investment is your hospital currently making in each of the following approaches to shift care to lower-acuity setting when appropriate? Other (please specify)	
A lot	12.5
Some	12.5
A little	0.0
None	75.0
N	24
B8_a. How much of an investment is your hospital currently making in each of the following approaches to limit medical overuse and/or not provide low-value services? Working with hospital-based providers to limit low-value services that occur within the hospital inpatient setting	
A lot	40.5
Some	40.5
A little	16.2
None	2.7
N	37
B8_b. How much of an investment is your hospital currently making in each of the following approaches to limit medical overuse and/or not provide low-value services? Partnering with non-hospital providers to limit low-value services that occur outside the hospital inpatient setting but have implications for follow-up services that occur in the hospital	
A lot	18.9
Some	45.9
A little	32.4
None	2.7
N	37
B8_c. How much of an investment is your hospital currently making in each of the following approaches to limit medical overuse and/or not provide low-value services? Enhancing palliative care options to limit intensive care at end of life that does not meet patient preferences	
A lot	51.4
Some	43.2
A little	2.7
None	2.7
N	37
B8_d. How much of an investment is your hospital currently making in each of the following approaches to limit medical overuse and/or not provide low-value services? Other (please specify)	
A lot	16.7
Some	5.6
A little	0.0
None	77.8
N	18

	Percent of hospitals
B10_a. How much of an investment is your hospital currently making in each of the following approaches to increase your hospital's global budget via the Market Shift Adjustment? Improving quality of services to attract new patients	
A lot	55.6
Some	22.2
A little	5.6
None	16.7
N	18
B10_b. How much of an investment is your hospital currently making in each of the following approaches to increase your hospital's global budget via the Market Shift Adjustment? Hiring new physicians or expanding clinics at the hospital	
A lot	38.9
Some	27.8
A little	16.7
None	16.7
N	18
B10_c. How much of an investment is your hospital currently making in each of the following approaches to increase your hospital's global budget via the Market Shift Adjustment? Opening new clinics in the community to increase referrals	
A lot	38.9
Some	22.2
A little	11.1
None	27.8
N	18
B10_d. How much of an investment is your hospital currently making in each of the following approaches to increase your hospital's global budget via the Market Shift Adjustment? Adding new lines of business (please specify)	
A lot	22.2
Some	33.3
A little	16.7
None	27.8
N	18
B10_e. How much of an investment is your hospital currently making in each of the following approaches to increase your hospital's global budget via the Market Shift Adjustment? Other (please specify)	
A lot	25.0
Some	0.0
A little	0.0
None	75.0
N	8

	Percent of hospitals
B13. Please indicate how much you agree or disagree with the following statement. Our hospital can operate in a financially viable way under the MD TCOC Model.	
Strongly agree	9.1
Agree	38.6
Neither agree nor disagree	13.6
Disagree	29.5
Strongly disagree	9.1
N	44
B14_a. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Patients' adherence to provider recommendations	
A great deal	25.0
Somewhat	68.2
A little	4.5
Not at all	2.3
N	44
B14_b. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Limited ongoing opportunities to reduce volume while maintaining high quality of care	
A great deal	25.0
Somewhat	34.1
A little	36.4
Not at all	4.5
N	44
B14_c. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Lack of access to other care options (e.g., urgent care or after hours at primary care)	
A great deal	18.2
Somewhat	47.7
A little	18.2
Not at all	15.9
N	44
B14_d. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Differences in payment and incentives between my hospital (global budgets) and other providers (fee for service)	
A great deal	54.5
Somewhat	20.5
A little	4.5
Not at all	20.5
N	44

	Percent of hospitals
B14_e. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Being at or near the cap for hospital price increases (from the rate corridors) due to past volume declines	
A great deal	20.9
Somewhat	37.2
A little	9.3
Not at all	32.6
N	43
B14_f. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Regulatory barriers, such as Medicare Inpatient Only list or providing supports to post-acute care providers (anti-kickback statutes)	
A great deal	23.8
Somewhat	45.2
A little	19.0
Not at all	11.9
N	42
B14_g. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Challenges operating during COVID-19 pandemic	
A great deal	81.4
Somewhat	4.7
A little	4.7
Not at all	9.3
N	43
B14_h. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Challenges operating under a fixed global budget with high inflation	
A great deal	86.0
Somewhat	4.7
A little	9.3
Not at all	0.0
N	43
B14_i. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019? Other (please specify)	
A great deal	41.7
Somewhat	4.2
A little	0.0
Not at all	54.2
N	24

Section C. Non-hospital spending

	Percent of hospitals
C1. How much influence do you think your hospital has on non-hospital spending among Medicare beneficiaries in your hospital's service area?	
A lot	0.0
Some	56.8
A little	34.1
None	9.1
N	44
C2. How much of an investment is your hospital currently making to limit growth in non-hospital Medicare spending?	
A lot	29.5
Some	36.4
A little	22.7
None	11.4
N	44
C3. How much has your hospital's investment in limiting growth in non-hospital Medicare spending changed since the 2019 implementation of the MD TCOC Model?	
Investment has increased	59.5
Investment has stayed the same	40.5
Investment has decreased	0.0
N	42
C4_a. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? The Medicare Performance Adjustment	
A lot	7.1
Some	45.2
A little	21.4
Not at all	26.2
N	42
C4_b. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Incentives for efficient episodes of care under ECIP	
A lot	2.4
Some	31.0
A little	40.5
Not at all	26.2
N	42
C4_c. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Incentives for efficient episodes of care under CTIs	
A lot	2.4
Some	42.9
A little	23.8
Not at all	31.0
N	42

	Percent of hospitals
C4_d. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Regional partnership grants to improve population health in ways that reduce need for non-hospital services	
A lot	11.9
Some	26.2
A little	26.2
Not at all	35.7
N	42
C4_e. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Payments for being part of a Care Transformation Organization supporting primary care in ways that reduce non-hospital spending	
A lot	31.0
Some	35.7
A little	19.0
Not at all	14.3
N	42
C4_f. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Accountable care organization contracts	
A lot	16.7
Some	14.3
A little	38.1
Not at all	31.0
N	42
C4_g. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending? Other (please specify)	
A lot	5.3
Some	0.0
A little	0.0
Not at all	94.7
N	19
C5_a. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Improving quality of hospital care to prevent post-discharge complications that may require additional non-hospital care	
A lot	81.0
Some	19.0
A little	0.0
None	0.0
N	42

	Percent of hospitals
C5_b. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Discharging patients to lower cost of care settings (for example, to home health instead of to skilled nursing facilities)	
A lot	44.2
Some	46.5
A little	7.0
None	2.3
N	43
C5_c. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Partnering with primary care or other providers to improve management of chronic conditions, reducing need for all types (hospital and non-hospital) of medical care	
A lot	65.1
Some	30.2
A little	4.7
None	0.0
N	43
C5_d. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Partnering with community-based organizations to address social determinants of health and prevent chronic conditions	
A lot	38.1
Some	19.0
A little	42.9
None	0.0
N	42
C5_e. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Partnering with primary care or other providers to increase adherence to clinical care guidelines (which could limit provision of low-value services)	
A lot	27.9
Some	41.9
A little	30.2
None	0.0
N	43
C5_f. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Partnering with primary care providers, specialty providers, or both to reduce duplication of imaging or testing	
A lot	9.3
Some	34.9
A little	51.2
None	4.7
N	43

	Percent of hospitals
C5_g. How much of an investment is your hospital currently making in each of the following approaches to limit growth in non-hospital spending? Other (please specify)	
A lot	0.0
Some	0.0
A little	0.0
None	100.0
N	18
C7_a. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Inability to affect patients you rarely or never interact with	
A lot	67.4
Some	27.9
A little	2.3
Not at all	2.3
N	43
C7_b. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Limited engagement of non-employed physicians	
A lot	20.9
Some	37.2
A little	34.9
Not at all	7.0
N	43
C7_c. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Limited incentives or supports for your hospital to make necessary investments	
A lot	23.3
Some	53.5
A little	14.0
Not at all	9.3
N	43
C7_d. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Difficulty partnering with outside organizations	
A lot	4.7
Some	58.1
A little	25.6
Not at all	11.6
N	43

	Percent of hospitals
C7_e. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Difficulty scaling effective interventions	
A lot	23.3
Some	58.1
A little	11.6
Not at all	7.0
N	43
C7_f. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Challenges operating during COVID-19 pandemic	
A lot	88.4
Some	7.0
A little	2.3
Not at all	2.3
N	43
C7_g. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Conflicting incentives (from global budgets) for hospitals to shift care outside of the hospital, which increases non-hospital spending	
A lot	46.5
Some	34.9
A little	16.3
Not at all	2.3
N	43
C7_h. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area? Other (please specify)	
A lot	13.6
Some	4.5
A little	0.0
Not at all	81.8
N	22

Section D. Participation in the Episode Care Improvement Program or Care Transformation Initiatives

	Percent of hospitals
D1_a. How important is each of the factors below in your hospital's participation in ECIP? Limited risk or downside to participating	
Very important	72.7
Somewhat important	22.7
Not very important	0.0
Not at all important	4.5
N	22
D1_b. How important is each of the factors below in your hospital's participation in ECIP? The opportunity for your hospital to earn incentive payments	
Very important	45.5
Somewhat important	45.5
Not very important	4.5
Not at all important	4.5
N	22
D1_c. How important is each of the factors below in your hospital's participation in ECIP? The ability for Care Partners to qualify for MACRA bonus payments	
Very important	63.6
Somewhat important	27.3
Not very important	4.5
Not at all important	4.5
N	22
D1_d. How important is each of the factors below in your hospital's participation in ECIP? The ability for your hospital to share incentives with care partners	
Very important	4.5
Somewhat important	36.4
Not very important	18.2
Not at all important	40.9
N	22
D1_e. How important is each of the factors below in your hospital's participation in ECIP? The ability for your hospital to share non-financial resources with care partners	
Very important	22.7
Somewhat important	31.8
Not very important	9.1
Not at all important	36.4
N	22

	Percent of hospitals
D1_f. How important is each of the factors below in your hospital's participation in ECIP? The chance to address an area of high spending or poor quality	
Very important	40.9
Somewhat important	22.7
Not very important	31.8
Not at all important	4.5
N	22
D1_g. How important is each of the factors below in your hospital's participation in ECIP? The interventions we deliver under ECIP should also help us perform well under global budgets (e.g., reduce preventable hospital use)	
Very important	13.6
Somewhat important	77.3
Not very important	4.5
Not at all important	4.5
N	22
D1_h. How important is each of the factors below in your hospital's participation in ECIP? Other (please specify)	
Very important	16.7
Somewhat important	0.0
Not very important	0.0
Not at all important	83.3
N	6
D2. In addition to—or instead of—financial incentives, are you providing any of the following non-financial incentives or supports to care partners as part of their participation in ECIP?	
Data	50.0
Don't know	4.8
Staff to support facility care partners (for example, care managers, embedded medical directors)	45.0
Technological assistance	10.0
Technology	20.0
We aren't providing any non-financial incentives or supports to ECIP care partners	45.0
N	20
D3_a. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Coordination and engagement with non-hospital providers, community organizations, and other partners	
A lot	4.5
Some	40.9
A little	13.6
Not at all	40.9
N	22

	Percent of hospitals
D3_b. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Challenges using data to drive and target changes	
A lot	9.1
Some	72.7
A little	9.1
Not at all	9.1
N	22
D3_c. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Low episode volume	
A lot	4.5
Some	13.6
A little	54.5
Not at all	27.3
N	22
D3_d. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Limited staff or bed capacity at post-acute care facilities	
A lot	27.3
Some	27.3
A little	0.0
Not at all	45.5
N	22
D3_e. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Administrative burden of participating in program	
A lot	19.0
Some	47.6
A little	23.8
Not at all	9.5
N	21
D3_f. To what extent are each of the following barriers that your hospital faces in its participation in ECIP? Other (please specify)	
A lot	0.0
Some	0.0
A little	0.0
Not at all	100.0
N	6

	Percent of hospitals
D4_a. How important is each of the factors below in your hospital's participation in CTIs? Avoid paying other hospitals via the Medicare Performance Adjustment – Efficiency Component	
Very important	23.1
Somewhat important	35.9
Not very important	30.8
Not at all important	10.3
N	39
D4_b. How important is each of the factors below in your hospital's participation in CTIs? The opportunity for the hospital to earn incentive payments	
Very important	51.3
Somewhat important	7.7
Not very important	38.5
Not at all important	2.6
N	39
D4_c. How important is each of the factors below in your hospital's participation in CTIs? The chance to be rewarded for past transformation success (i.e., transformation work that predates 2019)	
Very important	38.5
Somewhat important	12.8
Not very important	35.9
Not at all important	12.8
N	39
D4_d. How important is each of the factors below in your hospital's participation in CTIs? The opportunity to continue the same work we began under ECIP with greater flexibility or fewer administrative burdens	
Very important	35.9
Somewhat important	33.3
Not very important	17.9
Not at all important	12.8
N	39
D4_e. How important is each of the factors below in your hospital's participation in CTIs? The chance to address an area of high service use and spending	
Very important	50.0
Somewhat important	23.7
Not very important	21.1
Not at all important	5.3
N	38
D4_f. How important is each of the factors below in your hospital's participation in CTIs? The interventions we deliver under CTIs should also help us perform well under global budgets (e.g., reduce preventable hospital use)	
Very important	53.8
Somewhat important	20.5
Not very important	20.5
Not at all important	5.1
N	39

	Percent of hospitals
D4_g. How important is each of the factors below in your hospital's participation in CTIs? Other (please specify)	
Very important	5.9
Somewhat important	0.0
Not very important	5.9
Not at all important	88.2
N	17
D5_a. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Coordination and engagement with non-hospital providers, community organizations, and other partners	
A lot	27.5
Some	37.5
A little	10.0
Not at all	25.0
N	40
D5_b. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Challenges using data to drive and target changes	
A lot	37.5
Some	47.5
A little	10.0
Not at all	5.0
N	40
D5_c. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Low episode volume	
A lot	15.0
Some	27.5
A little	25.0
Not at all	32.5
N	40
D5_d. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Limited staff or bed capacity at post-acute care facilities	
A lot	42.1
Some	26.3
A little	2.6
Not at all	28.9
N	38
D5_e. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Administrative burden of participating in program	
A lot	18.4
Some	52.6
A little	21.1
Not at all	7.9
N	38

	Percent of hospitals
D5_f. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Regulations (for example, anti-kickback statutes) that limit or prevent sharing of incentives with providers not employed by our hospital or health system	
A lot	12.5
Some	12.5
A little	27.5
Not at all	47.5
N	40
D5_g. To what extent are each of the following barriers that your hospital faces in its participation in CTIs? Other (please specify)	
A lot	0.0
Some	5.3
A little	0.0
Not at all	94.7
N	19

Section E. Quality of care and the MD TCOC Model

	Percent of hospitals
E2_a. The Statewide Integrated Health Improvement Strategy (SIHIS) aligns three population health goals in the areas of obesity and diabetes, overdose mortality, and maternal and child health. To what extent is your hospital involved in efforts to help achieve each of these state-wide goals? Reduce mean body mass index (BMI) for adult residents	
A lot	29.3
Some	46.3
A little	24.4
Not at all	0.0
N	41
E2_b. The Statewide Integrated Health Improvement Strategy (SIHIS) aligns three population health goals in the areas of obesity and diabetes, overdose mortality, and maternal and child health. To what extent is your hospital involved in efforts to help achieve each of these state-wide goals? Improve overdose mortality	
A lot	65.9
Some	19.5
A little	9.8
Not at all	4.9
N	41
E6. To what extent is your hospital engaged in efforts to improve equity in health outcomes?	
A lot	68.2
Some	27.3
A little	2.3
Not at all	2.3
N	44
E7. To what extent are the actions your hospital is taking to improve equity in health outcomes encouraged by incentives or supports provided by the MD TCOC Model?	
A lot	14.0
Some	34.9
A little	25.6
Not at all	25.6
N	43

B.3.2d. Survey instrument

Evaluation of the Maryland Total Cost of Care Model

Hospital Survey

July 2022

Hospital Name
Hospital Address
City, State, Zip

INTRODUCTION

Mathematica was contracted by the Centers for Medicare & Medicaid Services (CMS) to conduct an independent evaluation of the Maryland Total Cost of Care (MD TCOC) Model. As part of that evaluation, we are seeking to collect information from all hospitals in Maryland to explore strategies used to improve care delivery in response to model incentives. We are genuinely interested in your observations on the way your hospital operates under the MD TCOC Model.

Your participation in the survey is voluntary but very important. Your responses will not have consequences for payments under the MD TCOC Model. Responses from your individual hospital will not be shared with anyone outside of CMS and the Mathematica research team. Mathematica will report the results of this survey to CMS in a nonidentifiable form that will ensure your confidentiality.

We estimate that the survey will take 45 minutes. The survey has five sections:

- Section A. Incentives and opportunities in the MD TCOC Model
- Section B. Hospital global budgets
- Section C. Non-hospital spending
- Section D. Participation in the Episode Care Improvement Program (ECIP) or Care Transformation Initiative (CTI)
- Section E. Quality of care and the MD TCOC Model

With help from HSCRC, we've identified you to complete the survey as a leading financial executive at your hospital. However, we encourage you to reach out to others at your hospital for further input. For example, you can ask others to answer questions or review responses as needed (e.g., the chief quality officer or other C-suite staff). The survey will be most accurate if it represents a consensus view of the hospital's key staff, arriving at the best answers after discussion. We understand that your hospital may be part of a system. For the purposes of this survey, please focus on the specific hospital listed at the top of this web page.

Geraldine Haile, Survey Director

MDTCOC@Mathematica.org

INSTRUCTIONS TO COMPLETE THE SURVEY

- ✓ If you are responsible for multiple hospitals, **please respond *only* about the hospital identified at the top of this web page.**
- ✓ Answer all questions to the best of your ability.
- ✓ If you answer “Other” for a question, please elaborate on your answer using the “*please specify*” line.
- ✓ For each item, please select only one answer unless instructions say to “SELECT ALL THAT APPLY.”

A. INCENTIVES AND OPPORTUNITIES IN THE MD TCOC MODEL

The MD TCOC Model includes a range of incentives and opportunities for hospitals. The model:

- **Continues all-payer global budgets, which began in [2014 if urban/2010 if rural] under the All-Payer Model.**
- **Adjusts hospital budgets based on a hospital's performance on quality measures and—via the Medicare Performance Adjustment (MPA)—total cost of care for attributed Medicare beneficiaries.**
- **Provides opportunities for hospitals to earn payments for efficient episodes of care (under the Episode Care Improvement Program [ECIP] or Care Transformation Initiative [CTI]), apply for new Regional Partnership Catalyst Grants, and function as Care Transformation Organizations (CTOs) for practices in the Maryland Primary Care Program (MDPCP).**

A1. To what extent have each of the following incentives and opportunities influenced the investments your hospital has made to change care delivery since the 2019 implementation of the MD TCOC Model?

Investment refers to contributions of staff, infrastructure, or other resources by your hospital.

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. All-payer hospital global budgets..................................	1 ❍	2 ❍	3 ❍	4 ❍
b. Quality adjustments to global budgets...........................	1 ❍	2 ❍	3 ❍	4 ❍
c. The Medicare Performance Adjustment	1 ❍	2 ❍	3 ❍	4 ❍
d. Incentives for efficient episodes of care under ECIP......	1 ❍	2 ❍	3 ❍	4 ❍
e. Incentives for efficient episodes of care under CTIs	1 ❍	2 ❍	3 ❍	4 ❍
f. Regional Partnership Grants	1 ❍	2 ❍	3 ❍	4 ❍
g. Payments for being part of a Care Transformation Organization under MDPCP	1 ❍	2 ❍	3 ❍	4 ❍

A2. Please indicate how much you agree or disagree with the following statement.

The incentives under the MD TCOC Model are clear, allowing our hospital to make informed decisions about how to respond to them.

1 ❍ Strongly agree
2 ❍ Agree
3 ❍ Neither agree nor disagree
4 ❍ Disagree
5 ❍ Strongly disagree

A3. [For hospitals that are part of a health system] To what extent are decisions about how your hospital responds to MD TCOC Model incentives made by your hospital's health system?

1 ❍ Fully – Decisions are centralized within the health system
2 ❍ Partly – System and hospital leadership jointly decide
3 ❍ Little – Leaders at our hospital largely make decisions on our own, with some input from the health system
4 ❍ None – Our hospital makes decisions independently, with no input from the health system

B. HOSPITAL GLOBAL BUDGETS

B1. How much of an investment is your hospital currently making in each of the following strategies to do well financially under your global budget?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Reducing preventable hospital use (that is, hospital use that is unplanned and may be prevented through improved care, care coordination, or effective community-based care)	1 ❍	2 ❍	3 ❍	4 ❍
b. Shifting care to lower-acuity settings (for example, ambulatory care settings) when appropriate	1 ❍	2 ❍	3 ❍	4 ❍
c. Limiting medical overuse, not providing low-value services, or both	1 ❍	2 ❍	3 ❍	4 ❍
d. Increasing your hospital's global budget via the Market Shift Adjustment..	1 ❍	2 ❍	3 ❍	4 ❍
e. Improving performance on quality measures that affect hospital budgets..	1 ❍	2 ❍	3 ❍	4 ❍
f. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

B2. How much has your hospital's investment in each of the following strategies changed <u>since the 2019 implementation of the MD TCOC Model?</u>

SELECT ONLY ONE PER ROW

	Investment has increased	Investment has stayed the same	Investment has decreased
a. Reducing preventable hospital use	1 ❍	2 ❍	3 ❍
b. Shifting care to lower-acuity settings when appropriate	1 ❍	2 ❍	3 ❍
c. Limiting medical overuse, not providing low-value services, or both	1 ❍	2 ❍	3 ❍
d. Increasing your hospital's global budget via the Market Shift Adjustment	1 ❍	2 ❍	3 ❍
e. Improving performance on quality measures that affect hospital budgets	1 ❍	2 ❍	3 ❍
f. Other (*please specify*) ________________________________	1 ❍	2 ❍	3 ❍

The questions on this page are about investments your hospital is currently making to *reduce preventable hospital use*. If you responded "a lot" or "some" to question B1.a, then please answer the questions on this page.

If your hospital is not making investments to *reduce preventable hospital use* (that is, you responded "a little" or "none" to question B1.a), please go to the next page.

B3. What type(s) of hospital use is your hospital primarily trying to reduce?

SELECT ALL THAT APPLY

1 ☐ Inpatient care

2 ☐ Outpatient care (not emergency department)

3 ☐ Emergency department care

B4. How much of an investment is your hospital currently making in each of the following approaches to *reduce preventable hospital use*?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Improving hospital care delivery (for example, discharge planning)	1 ❍	2 ❍	3 ❍	4 ❍
b. Coordinating medical or behavioral health care with non-hospital providers (for example, primary care, post-acute care)	1 ❍	2 ❍	3 ❍	4 ❍
c. Addressing social determinants of health using hospital staff	1 ❍	2 ❍	3 ❍	4 ❍
d. Addressing social determinants of health by coordinating with non-hospital providers or community-based organizations	1 ❍	2 ❍	3 ❍	4 ❍
e. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

B5. We're interested in learning about the specific investments and care delivery changes that your hospital is making to *reduce preventable hospital use*. Tell us more here.

The questions on this page are about investments your hospital is currently making to *shift care to lower-acuity settings when appropriate*. If you responded "a lot" or "some" to question B1.b, then please answer the questions on this page.

If your hospital is not making investments to *shift care to lower-acuity settings when appropriate* (that is, you responded "a little" or "none" to question B1.b), please go to the next page.

B6. How much of an investment is your hospital currently making in each of the following approaches to *shift care to lower-acuity setting when appropriate*?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Shifting emergency department care to non-hospital settings (for example, primary care, urgent care)..........	1 ❍	2 ❍	3 ❍	4 ❍
b. Shifting surgeries or other planned medical care to non-hospital settings (for example, ambulatory surgical centers, dialysis clinics)	1 ❍	2 ❍	3 ❍	4 ❍
c. Shifting hospitalized patients to skilled nursing facilities or other post-acute care more quickly (that is, reducing length of stay), when appropriate	1 ❍	2 ❍	3 ❍	4 ❍
d. Shifting care within your hospital to higher versus lower intensity services, when appropriate (for example, out of the intensive care unit)	1 ❍	2 ❍	3 ❍	4 ❍
e. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

B7. We're interested in learning about the specific investments and care delivery changes that your hospital is making to *shift care to lower-acuity settings when appropriate*. Tell us more here.

The questions on this page are about investments your hospital is currently making to *limit medical overuse and/or not provide low-value services*. If you responded "a lot" or "some" to question B1.c, then please answer the questions on this page.

If your hospital is not making investments *to limit medical overuse and/or not provide low-value services* (that is, you responded "a little" or "none" to question B1.c), please go to the next page.

B8. How much of an investment is your hospital currently making in each of the following approaches *to limit medical overuse and/or not provide low-value services?*

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Working with hospital-based providers to limit low-value services that occur within the hospital inpatient setting..	1 ❍	2 ❍	3 ❍	4 ❍
b. Partnering with non-hospital providers to limit low-value services that occur outside the hospital inpatient setting but have implications for follow-up services that occur in the hospital..	1 ❍	2 ❍	3 ❍	4 ❍
c. Enhancing palliative care options to limit intensive care at end of life that does not meet patient preferences	1 ❍	2 ❍	3 ❍	4 ❍
d. Other (*please specify*) ______________________	1 ❍	2 ❍	3 ❍	4 ❍

B9. We're interested in learning about the specific investments and care delivery changes your hospital is making to *limit medical overuse and/or not provide low-value services*. Tell us more here.

The questions on this page are about investments your hospital is currently making to *increase your hospital's global budget via the Market Shift Adjustment*. If you responded "a lot" or "some" to question B1.d, then please answer the questions on this page.

If your hospital is not making investments to *increase your hospital global budget via the Market Shift Adjustment* (that is, you responded "a little" or "none" to question B1.d), please go to the next page.

B10. How much of an investment is your hospital currently making in each of the following approaches to *increase your hospital's global budget via the Market Shift Adjustment*?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Improving quality of services to attract new patients......	1 ❍	2 ❍	3 ❍	4 ❍
b. Hiring new physicians or expanding clinics at the hospital	1 ❍	2 ❍	3 ❍	4 ❍
c. Opening new clinics in the community to increase referrals..	1 ❍	2 ❍	3 ❍	4 ❍
d. Adding new lines of business (*please specify*) ______________________________	1 ❍	2 ❍	3 ❍	4 ❍
e. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

B11. We're interested in learning about the specific care delivery changes that represent the most important investments your hospital is making to *increase your hospital's global budget via the Market Shift Adjustment*. Tell us more here.

The questions on this page are about investments your hospital is currently making to *improve performance on quality measures that affect hospital budgets.* If you responded “a lot” or “some” to question B1.e, then please answer the question B12.

If your hospital is <u>not</u> making investments to *improve performance on quality measures that affect hospital budgets* (that is, you responded “a little” or “none” to question B1.e), please go to question B13.

B12. We’re interested in learning about the specific investments and care delivery changes that your hospital is making to *improve performance on quality measures that affect hospital budgets*. Tell us more here.

B13. Please indicate how much you agree or disagree with the following statement.

Our hospital can operate in a financially viable way under the MD TCOC Model.

1 ❍ Strongly agree
2 ❍ Agree
3 ❍ Neither agree nor disagree
4 ❍ Disagree
5 ❍ Strongly disagree

B14. How much of a barrier is each of the following factors in your hospital's ongoing efforts to succeed under global budgets since the implementation of the MD TCOC Model in 2019?

SELECT ONLY ONE PER ROW

	A great deal	Somewhat	A little	Not at all
a. Patients' adherence to provider recommendations	1 ❍	2 ❍	3 ❍	4 ❍
b. Limited ongoing opportunities to reduce volume while maintaining high quality of care	1 ❍	2 ❍	3 ❍	4 ❍
c. Lack of access to other care options (e.g., urgent care or after hours at primary care)	1 ❍	2 ❍	3 ❍	4 ❍
d. Differences in payment and incentives between my hospital (global budgets) and other providers (fee for service)	1 ❍	2 ❍	3 ❍	4 ❍
e. Being at or near the cap for hospital price increases (from the rate corridors) due to past volume declines	1 ❍	2 ❍	3 ❍	4 ❍
f. Regulatory barriers, such as Medicare Inpatient Only list or providing supports to post-acute care providers (anti-kickback statutes)	1 ❍	2 ❍	3 ❍	4 ❍
g. Challenges operating during COVID-19 pandemic	1 ❍	2 ❍	3 ❍	4 ❍
h. Challenges operating under a fixed global budget with high inflation..	1 ❍	2 ❍	3 ❍	4 ❍
i. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

C. NON-HOSPITAL SPENDING

Under the MD TCOC Model, the state of Maryland is now responsible for limiting the growth of total Medicare Part A and B spending, not just hospital facility spending. CMS and Maryland have introduced new hospital incentives designed to engage hospitals in limiting the growth in total Medicare spending, including non-hospital spending.

In this section, we are interested in learning more about your perceptions of and efforts toward (if any) controlling spending that falls outside of hospitals' global budgets.

C1. How much influence do you think your hospital has on non-hospital spending among Medicare beneficiaries in your hospital's service area?

1 ❍ A lot
2 ❍ Some
3 ❍ A little
4 ❍ None

C2. How much of an investment is your hospital currently making to limit growth in non-hospital Medicare spending?

1 ❍ A lot
2 ❍ Some
3 ❍ A little
4 ❍ None

C3. How much has your hospital's investment in limiting growth in non-hospital Medicare spending changed since the 2019 implementation of the MD TCOC Model?

1 ❍ Investment has increased
2 ❍ Investment has stayed the same
3 ❍ Investment has decreased

C4. To what extent have the following incentives encouraged your hospital to limit growth in non-hospital Medicare spending?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. The Medicare Performance Adjustment	1 ❍	2 ❍	3 ❍	4 ❍
b. Incentives for efficient episodes of care under ECIP......	1 ❍	2 ❍	3 ❍	4 ❍
c. Incentives for efficient episodes of care under CTIs	1 ❍	2 ❍	3 ❍	4 ❍
d. Regional partnership grants to improve population health in ways that reduce need for non-hospital services...	1 ❍	2 ❍	3 ❍	4 ❍
e. Payments for being part of a Care Transformation Organization supporting primary care in ways that reduce non-hospital spending	1 ❍	2 ❍	3 ❍	4 ❍
f. Accountable care organization contracts	1 ❍	2 ❍	3 ❍	4 ❍
g. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

C5. How much of an investment is your hospital currently making in each of the following approaches *to limit growth in non-hospital spending*?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	None
a. Improving quality of hospital care to prevent post-discharge complications that may require additional non-hospital care	1 ❍	2 ❍	3 ❍	4 ❍
b. Discharging patients to lower cost of care settings (for example, to home health instead of to skilled nursing facilities)...	1 ❍	2 ❍	3 ❍	4 ❍
c. Partnering with primary care or other providers to improve management of chronic conditions, reducing need for all types (hospital and non-hospital) of medical care..	1 ❍	2 ❍	3 ❍	4 ❍
d. Partnering with community-based organizations to address social determinants of health and prevent chronic conditions..	1 ❍	2 ❍	3 ❍	4 ❍
e. Partnering with primary care or other providers to increase adherence to clinical care guidelines (which could limit provision of low-value services)....................	1 ❍	2 ❍	3 ❍	4 ❍
f. Partnering with primary care providers, specialty providers, or both to reduce duplication of imaging or testing	1 ❍	2 ❍	3 ❍	4 ❍
g. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

Question C6 is for hospitals that are making an effort to *limit growth in non-hospital spending*. If you responded "a lot" or "some" to any sub-item in question C5, please respond to this question. If you answered "a little" or "none" to all sub-items in question C6, please skip to question C7.

C6. We're interested in learning more about the specific care delivery changes your hospital is making to *limit growth in non-hospital spending*. Tell us more here.

C7. To what extent are each of the following barriers that your hospital faces in limiting non-hospital spending for Medicare beneficiaries in your hospital's service area?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. Inability to affect patients you rarely or never interact with	1 ❍	2 ❍	3 ❍	4 ❍
b. Limited engagement of non-employed physicians	1 ❍	2 ❍	3 ❍	4 ❍
c. Limited incentives or supports for your hospital to make necessary investments	1 ❍	2 ❍	3 ❍	4 ❍
d. Difficulty partnering with outside organizations	1 ❍	2 ❍	3 ❍	4 ❍
e. Difficulty scaling effective interventions	1 ❍	2 ❍	3 ❍	4 ❍
f. Challenges operating during COVID-19 pandemic	1 ❍	2 ❍	3 ❍	4 ❍
g. Conflicting incentives (from global budgets) for hospitals to shift care outside of the hospital, which increases non-hospital spending	1 ❍	2 ❍	3 ❍	4 ❍
h. Other (*please specify*) ______________________	1 ❍	2 ❍	3 ❍	4 ❍

D. PARTICIPATION IN ECIP OR CTI

The questions in this section are for hospitals that are participating in ECIP, CTI, or both. If your hospital is not participating in ECIP or CTI, please skip to Section E.

Questions D1 to D3 are for hospitals participating in ECIP. If your hospital is participating in CTI only, skip to question D4.

The Episode Care Improvement Program (ECIP), which began in 2019, pays hospitals for successfully working with non-hospital partners to reduce total costs for episodes of care that start in the hospital but end 90 days later. Under ECIP, 5 percent of the reconciliation payment is contingent on a hospital's performance on a composite quality score.

D1. How important is each of the factors below in your hospital's participation in ECIP?

SELECT ONLY ONE PER ROW

	Very important	Somewhat important	Not very important	Not at all important
a. Limited risk or downside to participating	1 ❍	2 ❍	3 ❍	4 ❍
b. The opportunity for your hospital to earn incentive payments	1 ❍	2 ❍	3 ❍	4 ❍
c. The ability for Care Partners to qualify for MACRA bonus payments	1 ❍	2 ❍	3 ❍	4 ❍
d. The ability for your hospital to share incentives with care partners	1 ❍	2 ❍	3 ❍	4 ❍
e. The ability for your hospital to share non-financial resources with care partners	1 ❍	2 ❍	3 ❍	4 ❍
f. The chance to address an area of high spending or poor quality	1 ❍	2 ❍	3 ❍	4 ❍
g. The interventions we deliver under ECIP should also help us perform well under global budgets (e.g., reduce preventable hospital use)	1 ❍	2 ❍	3 ❍	4 ❍
h. Other (*please specify*) ________________	1 ❍	2 ❍	3 ❍	4 ❍

D2. In addition to—or instead of—financial incentives, are you providing any of the following non-financial incentives or supports to care partners as part of their participation in ECIP?

SELECT ALL THAT APPLY

1 ☐ Staff to support facility care partners (for example, care managers, embedded medical directors)
2 ☐ Data
3 ☐ Technology
4 ☐ Technological assistance
5 ☐ Other *(please specify)* ______________________
6 ❍ Don't know
0 ❍ We aren't providing any non-financial incentives or supports to ECIP care partners

D3. To what extent are each of the following barriers that your hospital faces in its participation in ECIP?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. Coordination and engagement with non-hospital providers, community organizations, and other partners	1 ❍	2 ❍	3 ❍	4 ❍
b. Challenges using data to drive and target changes	1 ❍	2 ❍	3 ❍	4 ❍
c. Low episode volume	1 ❍	2 ❍	3 ❍	4 ❍
d. Limited staff or bed capacity at post-acute care facilities	1 ❍	2 ❍	3 ❍	4 ❍
e. Administrative burden of participating in program	1 ❍	2 ❍	3 ❍	4 ❍
f. Other (*please specify*) ______________	1 ❍	2 ❍	3 ❍	4 ❍

Questions D4 to D5 are for hospitals participating in a CTI. If your hospital is not participating in a CTI, skip to Section E.

Care Transformation Initiatives (CTIs) pay hospitals for efficient episodes of care and allow hospitals flexibility in defining the episodes and interventions.

D4. How important is each of the factors below in your hospital's participation in CTIs?

SELECT ONLY ONE PER ROW

	Very important	Somewhat important	Not very important	Not at all important
a. Avoid paying other hospitals via the Medicare Performance Adjustment – Efficiency Component	1 ❍	2 ❍	3 ❍	4 ❍
b. The opportunity for the hospital to earn incentive payments	1 ❍	2 ❍	3 ❍	4 ❍
c. The chance to be rewarded for past transformation success (i.e., transformation work that predates 2019).....	1 ❍	2 ❍	3 ❍	4 ❍
d. The opportunity to continue the same work we began under ECIP with greater flexibility or fewer administrative burdens	1 ❍	2 ❍	3 ❍	4 ❍
e. The chance to address an area of high service use and spending..	1 ❍	2 ❍	3 ❍	4 ❍
f. The interventions we deliver under CTIs should also help us perform well under global budgets (e.g., reduce preventable hospital use)...	1 ❍	2 ❍	3 ❍	4 ❍
g. Other (*please specify*) ________________________	1 ❍	2 ❍	3 ❍	4 ❍

D5. To what extent are each of the following barriers that your hospital faces in its participation in CTIs?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. Coordination and engagement with non-hospital providers, community organizations, and other partners ...	1 ❍	2 ❍	3 ❍	4 ❍
b. Challenges using data to drive and target changes	1 ❍	2 ❍	3 ❍	4 ❍
c. Low episode volume	1 ❍	2 ❍	3 ❍	4 ❍
d. Limited staff or bed capacity at post-acute care facilities	1 ❍	2 ❍	3 ❍	4 ❍
e. Administrative burden of participating in program	1 ❍	2 ❍	3 ❍	4 ❍
f. Regulations (for example, anti-kickback statutes) that limit or prevent sharing of incentives with providers not employed by our hospital or health system	1 ❍	2 ❍	3 ❍	4 ❍
g. Other (*please specify*) __________________________	1 ❍	2 ❍	3 ❍	4 ❍

E. QUALITY OF CARE AND MD TCOC MODEL

E1. Please describe any unintended consequences that have resulted from the MD TCOC Model. Consequences could refer to changes in quality, access, finance or other domains.

E2. The Statewide Integrated Health Improvement Strategy (SIHIS) aligns three population health goals in the areas of obesity and diabetes, overdose mortality, and maternal and child health. To what extent is your hospital involved in efforts to help achieve each of these state-wide goals?

SELECT ONLY ONE PER ROW

	A lot	Some	A little	Not at all
a. Reduce mean body mass index (BMI) for adult residents	1 ❍	2 ❍	3 ❍	4 ❍
b. Improve overdose mortality	1 ❍	2 ❍	3 ❍	4 ❍

E3. [For hospitals that answered a lot or some on E2.a] What is your hospital doing to help achieve the goal of reducing mean BMI for adult residents?

E4. [For hospitals that answered a lot or some on E2.b] What is your doing to help achieve the goal of improving overdose mortality?

E5. **How, if at all, did the COVID-19 pandemic impact the investments your hospital was able to make under the MD TCOC Model?**

E6. **To what extent is your hospital engaged in efforts to improve equity in health outcomes?**

1 ❍ A lot
2 ❍ Some
3 ❍ A little
4 ❍ Not at all

E6a. **[For hospitals that answered a lot or some on E6] Please tell us more about what your hospital is doing to improve equity in health outcomes.**

E7. **To what extent are the actions your hospital is taking to improve equity in health outcomes encouraged by incentives or supports provided by the MD TCOC Model?**

1 ❍ A lot
2 ❍ Some
3 ❍ A little
4 ❍ Not at all

Thank you for taking the time to complete this survey.

Appendix C. Methods and Supplemental Results for Trends in Medicaid and Children's Health Insurance Program Outcomes

This appendix describes how we implemented regression-adjusted trend analyses for the Maryland Medicaid and Children's Health Insurance Program (CHIP) population and the Maryland Medicare fee-for-service (FFS) population from 2014 to 2021, the most recent year available for Medicaid and CHIP data. We first describe our approach to the regression-adjusted trend analyses (Section C.1) and then describe how we defined the outcomes measures, including any workarounds needed for the Medicaid data (Section C.2). We subsequently describe how we constructed the analysis files, including covariates used in the regressions (Section C.3). We conclude by reporting supplemental results; these include a detailed table showing how the Medicaid population changed from 2014 to 2021 and additional trend analyses with separate trends lines for dually eligible versus non-dually eligible Medicare FFS beneficiaries (Section C.4).

C.1. Design and methods for estimating Medicaid trends

C.1.1. Medicaid data sources

There are two possible federal data sources for Maryland Medicaid and CHIP enrollment and claims and encounter data from 2011 to 2021. The most recent data source, Transformed Medicaid Statistical Information System (T-MSIS) Analytic File (TAF), is available for Maryland Medicaid and CHIP for 2014 to 2021. For 2011 to 2013, the relevant data source is the Medicaid Analytic eXtract (MAX).

C.1.2. Medicaid study design

The goal of these analyses was to estimate changes in outcomes for Maryland's Medicaid population during the Maryland Model. The Medicaid trend analysis included only Maryland Medicaid and CHIP enrollees and only from 2014 to 2021, the most recent year of TAF data available. We limited the analyses to Maryland Medicaid and CHIP enrollees only because of the challenges in finding comparison states that (1) expanded Medicaid at the same time as Maryland (January 2014); (2) switched from reporting MAX to TAF at the same time as Maryland (January 2014); and (3) have similar Medicaid programs, policies, and populations. We limited the analyses to 2014 and later because we expected the quality and completeness of the MAX and TAF to differ. Because Maryland switched from MAX to TAF concurrently with the launch of the Maryland All-Payer Model in 2014, we would be unable to interpret whether any changes in trends were because of changes in data quality or the Maryland Model.

C.1.3. Methods for estimating trends

Trends in outcomes without a comparison group could change for a variety of reasons, including because the underlying population has changed—as we expect it would have in Maryland during this time period with the expansion of Medicaid. To account for this, we used regression methods that control for changes in the composition of the Medicaid population over time.

Specifically, we calculated regression-adjusted trends for each year in the analyses using the distribution of covariates among the population in 2014. By holding the characteristics of the population constant, changes in the characteristics of the population will not drive trends. This is important because the number and percentage of enrollees in the adult expansion population grew over time. In addition, in response to the COVID-19 pandemic in March 2020, Congress increased Medicaid funding to states. In exchange, states were not allowed to terminate enrollees' Medicaid coverage until the public health emergency (PHE) period ended, unless enrollees requested disenrollment, moved out of state, or died (Erzouki 2022; Wikle and Wagner 2022). As a result of this continuous enrollment requirement, the composition of the Medicaid population in our analyses during the PHE period (2020 and 2021) likely differed from the population in the period before the PHE (2014 to 2019) because beneficiaries who would have lost coverage in the absence of the PHE remained in our analysis.

To estimate regression-adjusted trends, we fit the following model for each outcome:

$$y_{it} = \gamma_t + X_{it}\beta + \gamma_t X_{it}\delta + \varepsilon_{it}$$

In this model, y_{it} represents outcome *y* for enrollee *i* in year *t;* γ_t represents a set of year dummy variables, one each for 2015 through 2022 (with 2014 as the reference year); X_{it} is the set of independent demographic and enrollment covariates whose relationship with the outcome is allowed to change over time via the interaction term, $\gamma_t * X_{it}$; and ε_{it} is the error term. Using the Maryland Medicaid population in 2014, we applied the coefficients from the regression model to calculate the adjusted mean value of each outcome in each year (holding the characteristics of the population fixed based on the characteristics of the 2014 population).

To facilitate direct comparison of trends with Medicare FFS, we fit similar regression models for our Medicare FFS populations over the same years, overall and stratified by dual eligibility status. We describe the outcomes and covariates included in the regressions in the following sections.

C.2. Measures, definitions, and Medicaid analytic file construction

To construct Medicaid claims-based outcomes at the beneficiary-year level, we relied on the Maryland TAF inpatient (IP), long-term (LT), other services (OT), and pharmacy (RX) files from the Virtual Research Data Center. These files contain FFS claims and managed care encounter records on all services funded by Medicaid. The OT file also contains records for monthly capitated payments made by the state for enrollees in managed care plans. As we describe in more detail below, we used these records, along with payments on FFS claims, to measure total spending. Our input Medicare analysis files, and all claims-based outcomes used in these trend analyses, are the same as those described in Appendix B. For this reason, the following sections focus primarily on the Medicaid measures, although we include a section describing how we inflated both total per-capita Medicare FFS spending and Medicaid spending to 2021 dollars for the trend analyses.

Maryland TAF data contained a few anomalies in the IP, LT, and OT files that required workarounds. Exhibit C.1 describes these workarounds and their implications for the claims-based outcomes measures and Chronic Conditions Data Warehouse (CCW) chronic condition variables used to describe characteristics of the population.

Exhibit C.1. Data issues and workarounds in the IP, LT, and OT files

TAF data issues	Workarounds
The IP, LT, and OT files have unusable diagnosis code data from October 2015 to December 2017[a]	• We analyzed and reported measures that depend on diagnosis codes (that is, potentially avoidable hospitalizations and asthma-related ED visits and observation stays for children ages 2 to 17) only in years with usable diagnosis data (2014 and 2018 to 2021). • We did not adjust for CCW chronic condition categories in our regression models. Instead, we compared the distribution of CCW chronic condition categories among the Medicaid and CHIP population in 2014 and 2021, as shown in Exhibit D.4, to ascertain how the health status of the population changed over time as a result of the adult expansion and the PHE period, and any other changes affecting Medicaid and CHIP enrollment over the analysis period. In general, we find similarities across time in the prevalence of many conditions, with the exception of several mental health conditions including anxiety disorders, depressive disorders, and post-traumatic stress disorder, which have higher prevalence in 2021 than in 2014.
The IP file has unusable procedure code data from 2014 to 2019[b]	• We measured potentially preventable hospitalizations for the adult Medicaid and CHIP population without applying any procedure code-based criteria. Practically, this means that, for adults ages 18 and older, we did not apply procedure-based exclusions to PQIs 07 (Hypertension), 08 (Heart failure), 11 (Community-acquired pneumonia), and 12 (Urinary tract infection) which could increase the rate of potentially avoidable hospitalizations. Similarly, we were unable to calculate PQI 16 (lower extremity amputation among patients with diabetes), which could decrease the rate of potentially avoidable hospitalizations. • For children ages 6 to 17, we did not apply procedure-based exclusions to Pediatric Quality Indicator 18 (Urinary tract infection), which could increase the rate of potentially avoidable hospitalizations.
Billing provider identifiers are missing on many IP claims until 2019	• To identify claims associated with the same inpatient hospitalization, we did not attempt to identify transfers between different facilities (which we do for the Medicare analyses). We simply identified inpatient hospital claims for the same enrollee with admission and discharge dates that overlapped or when the admission date on a subsequent claim was the same day or one day after the discharge date on a prior claim. Compared with Medicare, this might collapse relatively more claims into a single stay because we effectively treat all claims with an admission date on the same day or following day of a prior discharge date as a transfer, whereas some might not be transfers.
Most managed care enrollees had missing capitated payment data in the January 2014 OT file	• We backfilled the January 2014 capitated payment amounts using the February 2014 capitated payment amounts for enrollees who were enrolled in managed care in both months and had missing data in January 2014.

Source: Mathematica's analyses of Maryland TAF data.

[a] The unusable rating comes from the TAF DQ Atlas (here for 2016 ratings, for example: https://www.medicaid.gov/dq-atlas/landing/states/single?state=24&tafVersionId=7). Our analyses of the Maryland TAF data confirm this rating.

[b] The unusable rating comes from the TAF DQ Atlas (here for 2018 ratings, for example: https://www.medicaid.gov/dq-atlas/landing/states/single?state=24&tafVersionId=34). Our analyses of the Maryland IP data confirm this rating.

CCW = Chronic Conditions Data Warehouse; DQ = Data Quality; ED = emergency department; IP = inpatient; LT = long-term; OT = other services; PQI = Prevention Quality Indicator; TAF = Transformed Medicaid Statistical Information System Analytic File.

We report annualized claims-based measures based on the total number of months that each enrollee was observable in the Medicaid claims and encounter data (that is, they were alive and enrolled exclusively in Medicaid or CHIP with full benefits). Exclusive enrollment in Medicaid or CHIP signifies that the enrollee was not dually eligible for Medicare and Medicaid (we include these enrollees in our Medicare FFS analyses) and did not have other third-party coverage.[34] For example, if a beneficiary was observable for 10 of 12 months of the year, and we observed one hospitalization in Medicaid IP file for this beneficiary over 10 months, then their annualized number of all-cause hospitalizations would be 1.2 (that is, the average number of hospitalizations per month—one divided by 10—multiplied by 12 months). Similar to what we do for Medicare, we then created observability weights that represent the proportion of total calendar time in a year that beneficiaries were alive and enrolled exclusively in Medicaid or CHIP with full benefits. Using the same example above, a beneficiary observable for 10 months with a single hospitalization would have an annualized number of hospitalizations of 1.2 and an observability weight of 0.83 (10/12). All regression analyses and reported trends were weighted by observability.

C.2.1. Medicaid hospital use measures

All-cause acute care hospital admissions (number of admissions per beneficiary per year)

This measure is the annualized number of hospitalizations reported in the TAF IP file for the beneficiary during the year. These include IP claims with a type of bill value indicating an inpatient hospital (BILL_TYPE_CD = 011x or 012x, where "x" is any digit) or where the billing provider taxonomy code indicated that the facility was any of the following: General Acute Care Hospital, General Acute Care Hospital – Children; General Acute Care Hospital – Critical Access; General Acute Care Hospital – Rural; General Acute Care Hospital – Women; Special Hospital; or Epilepsy Unit. Multiple claims for acute admissions with overlapping or contiguous dates were combined into a single record. Consistent with the Medicare all-cause hospitalization measure, we excluded hospitalizations for psychiatric care,[35] inpatient rehabilitation stays, and long-term hospital stays. We also excluded hospitalizations for pregnancy, delivery, or neonates from our measure of all-cause hospitalizations using publicly available specifications and code sets (MACBIS 2023; Harrison et al. 2023). Depending on the year, 50 to 60% of all hospitalizations in Maryland were for pregnancy, delivery, and neonates. We discovered, however, that the algorithm to assign unique CCW beneficiary identifiers contained errors in 2020 and 2021 in Maryland. Specifically, infants with the same birth dates in 2020 and 2021 all had the same unique CCW beneficiary identifier, even though they had different state Medicaid identifiers, different claim identifiers, and oftentimes different diagnosis codes, different procedure codes, and different billing provider IDs. Including these hospitalizations in our measure of all-cause hospitalizations in 2020 and 2021 would artificially lower the rate of hospitalizations because we count multiple claims for the same beneficiary with the same or overlapping dates as one hospitalization.

[34] We also excluded a small number of additional enrollees in each year who had missing date of birth data.

[35] Diagnosis data were unusable from October 2015 to December 2017. For this reason, we tested an alternative measure of all-cause hospitalizations that included stays for psychiatric conditions to assess whether potential misclassification of psychiatric stays (defined using principal diagnosis on inpatient claims) during this period might have skewed trends. Unadjusted and adjusted trends were materially the same for all-cause hospitalizations with or without psychiatric stays (suggesting that diagnosis data were more reliable for mental health conditions in all years). For this reason, we reported the measure without psychiatric stays for consistency with Medicare analyses.

Outpatient ED visits and observation stays (number of visits per beneficiary per year)

This measure is the annualized number of outpatient ED visits and observation stays for the beneficiary during the year that do not lead to a hospitalization. Visits that do not lead to a hospitalization are identified in the OT file using both facility and professional claims. For ED visits, we flagged facility claims as outpatient ED visits if they contained revenue center line items equal to 045X (where "X" can take any value) or 0981 (emergency room care). We flagged professional claims as ED visits if they contained an ED-related evaluation and management services (99281 – 99285) or if the claim place of service code indicated that care was delivered in the ED and the claim included procedure codes, as long as at least one procedure code on the claim was not for laboratory and radiology services (that is, we excluded ED visits that might have been planned visits for pre-admission testing). We identified outpatient observation stays on facility claims using the same definition as used for Medicare analyses (that is, with revenue center codes or 0760 or 0762 and procedure code of G0378 and number of hours equal to eight or more). On professional claims, we flagged observation visits if they had an observation-specific professional procedure code (99217-99220; 99224-99226; 99234-99236). We then cross-checked all OT claims flagged as ED visits or observation stays against the IP claims and excluded any claims from our definition of outpatient ED visits and observation stays if (1) we found an inpatient claim for the same beneficiary and (2) the ED or observation service dates fell anytime on or between the inpatient admission and discharge dates. Consistent with Medicare analyses, we then capped the number of either type of visit (observation stays and ED visits) to one per day.

Asthma-related ED visits among children ages 2 to 17 (number of visits per beneficiary per year)

We flagged outpatient ED visits and observation stays as asthma-related for children ages 2 to 17 if the beneficiary was age 2 to 17 on the discharge date of the claim and the claim contained an asthma diagnosis in the principal diagnosis field. In the ICD-9 period, these codes included: 49300, 49301, 49302, 49310, 49311, 49312, 49320, 49321, 49322, 49381, 49382, 49390, 49391, or 49392. In the ICD-10 period, these codes included J4520, J4521, J4522, J4530, J4531, J4532, J4540, J4541, J4542, J4550, J4551, J4552, J45901, J45902, J45909, J45990, J45991, or J45998.

C.2.2. Quality of care measures

Potentially preventable admissions (number of admissions per beneficiary per year)

This outcome is measured separately for adults and children. Among Medicaid adults ages 18 and older, this measure reflects the annualized number of hospitalizations for the enrollee during the year in which the admission met the criteria for the Prevention Quality Indicators (PQI) overall composite measure (PQI #90). To construct this measure, we applied the Agency for Healthcare Research and Quality's 2020 Quality Indicators Software to all inpatient hospital claims (defined earlier) and then counted the number of hospital admissions for the enrollee each year that the software flagged as being an admission for one of the following PQIs: diabetes short-term complications (PQI #01), diabetes long-term complications (PQI #03), chronic obstructive pulmonary disease or asthma in older adults (PQI #05), hypertension (PQI #07), heart failure (PQI #08), community-acquired pneumonia (PQI #11), urinary tract infection (PQI #12), uncontrolled diabetes (PQI #14), or asthma in younger and adults (PQI #15) (AHRQ n.d.[a]). As described in Exhibit C.1, we did not apply procedure-based exclusion codes to the PQIs because of poor quality procedure code data in Maryland's IP file in most years. We also did not measure the PQI for lower extremity amputation among patients with diabetes (PQI #16) in our measure of potentially avoidable hospitalizations for the same reason.

For children ages 6 to 17, this measure reflects the annualized number of hospitalizations for the enrollee during the year in which the admission met the criteria for the Pediatric Quality Indicators (PDI) overall composite measure (PQI #90). To construct this measure, we applied the Agency for Healthcare Research and Quality's 2020 Pediatric Quality Indicators Software to all inpatient hospital claims (defined earlier) and then counted the number of hospital admissions for the child each year that the software flagged as being admissions for one of the following PDIs: asthma (PDI #14), diabetes short-term complications (PDI #15), gastroenteritis (PDI #16), or urinary tract infection (PDI #18) (AHRQ n.d.[b]). Consistent with our approach to measuring PQIs for adults, we did not apply the procedure-based exclusion codes to the PDI for urinary tract infection (the only PDI that used procedure code exclusions) because of poor quality procedure code data in Maryland's IP file in most years.

C.2.3. Total Medicaid per-capita spending per year

Total Medicaid per-capita spending per year

We measured total Medicaid spending per enrollee per year as the sum of the following categories of spending in the year: (1) total monthly capitated payments made by Maryland Medicaid to participating Medicaid managed care organizations for managed care enrollees (CLM_TYPE_CD = 2, B, V), (2) total FFS spending for services carved out of managed care contracts (CLM_TYPE_CD = 1, A, U), (3) supplemental payments made to providers (for example, for services paid above the capitation rate or above a negotiated rate) (CLM_TYPE_CD = 5, E, Y), and (4) service tracking claims that reflect lump sum payments to providers (for example, to disproportionate share hospitals) (CLM_TYPE_CD = 4, D, X). We mostly found records for service tracking claims in 2020,[36] which might reflect additional payments made to hospitals because of the PHE. The lump-sum payments made to providers on service tracking claims cannot be linked to individual enrollees. Thus, for this category of spending, we summed all such payments in each year (when relevant) and spread the dollars across all enrollees who were ever observable in the TAF in the year.

Total Medicare FFS and Medicaid inflation-adjusted per-capita spending per year

For the trend analyses, we inflated 2014 to 2020 spending amounts to 2021 dollars for both Medicare FFS and Medicaid analyses. We used the Gross Domestic Product price inflator, which is recommended for health care trend analyses. The Gross Domestic Product price inflator reflects federal and non-federal spending across a broad swath of the U.S. economy (AHRQ n.d.[c]).

C.2.4. Medicaid analytic file construction

We constructed the Medicaid trend analysis file at the enrollee-year level. There is one observation per enrollee in each year since 2014 that they were enrolled in Medicaid or CHIP. People enrolled in Medicaid or CHIP in multiple years will have multiple observations in the file, and those enrolled for only one year will have only one observation. We limited the analysis file to those who were observable for at least one month in the claims or encounter data during the year (that is, they were alive, not dually enrolled in Medicare and Medicaid, eligible for full benefits, and did not have any other third-party coverage).

Our approach to developing the analysis file started with an assessment of TAF data quality and completeness to identify workarounds needed and to confirm we can consistently construct all the outcomes and regression covariates needed for our study in all TAF years. In addition to the workarounds identified and implemented for the IP, LT, OT, and RX files as described in Exhibit C.1, Exhibit C.2 describes several data issues and workarounds implemented for the annual demographic and eligibility (DE) file.

[36] One possible explanation is that, per the TAF claims data dictionary (https://www2.ccwdata.org/web/guest/data-dictionaries), “RIFs prior to August 2021 did not include these service tracking claims.”

Exhibit C.2. Data issues and workarounds required for Maryland's annual DE file

TAF data issue	Workaround
TAF DQ Atlas rated the quality of the dual eligibility variable as high concern in 2015	We flagged enrollees as dually enrolled in Medicare and Medicaid in any month if the TAF DE file indicated dual enrollment in the month or the Medicare Beneficiary Summary File indicated dual enrollment in the month (there was high concordance between the two data sources in most years).
TAF DQ Atlas rated the quality of the CHIP enrollment data as medium or high concern between 2014 and 2016	We did not adjust for whether child enrollees were Medicaid or CHIP enrollees in regression analyses.
All enrollees had missing eligibility data for March 2015	We backfilled the March 2015 eligibility data for those enrolled in either February or April 2015, depending on enrollment in those and other months of the year, evidence of health care use in the claims or encounter data in March 2015; evidence of capitated payments made by the state for the enrollee in March 2015; and date of birth.

Source: Mathematica's analyses of Maryland TAF data.

CHIP = Children's Health Insurance Program; DE = demographic and eligibility file; DQ = Data Quality; TAF = Transformed Medicaid Statistical Information System Analytic File.

The construction of the analytic file involved the following steps:

1. Pulling enrollment and demographic information for the full Medicaid and CHIP population (that is, all people who were ever enrolled in Medicaid or CHIP) during each year from the TAF DE file
2. Identifying the first observable month, if any, among the full Medicaid and CHIP population (many enrollees are never observable during the year because, for example, they are dually enrolled in Medicare and Medicaid for the full year) for the claims-based analyses
3. Constructing variables to reflect demographic and enrollment characteristics for the year, as described in Exhibit D.3; characteristics that could change within any year, such as major eligibility category, are characterized based on enrollees' data in the first observable month of the year
4. Developing claims-based measures over all observable months for all enrollees and merging these measures to the enrollee-level file by unique beneficiary identifier
5. Annualizing the claims-based measures based on the number of months observable
6. Inflating spending outcomes measures between 2014 and 2020 to 2021 dollars
7. Merging on data from external sources (this included merging American Community Survey data on zip code-level rurality and race and ethnicity composition as well as merging on the mean SVI summary score for each zip code using all zip codes in our Medicare analysis files)
8. Applying a final set of exclusion criteria for each year's file to exclude enrollees from the analytic sample if they were never observable (for example, dually enrolled in Medicare and Medicaid in all months) or had missing or inaccurate zip code data.

We also constructed CCW-like condition categories in 2014 and 2021. We describe these as CCW-like because we constructed them using claims and encounter data within each year only (that is, we did not look back into prior years' claims and encounter data, as typically required by the algorithms for many of the CCW conditions). We limited the construction of these variables to a single year for three reasons. First, we would have needed to combine MAX and TAF data to construct the CCW variables for 2014, but we had concerns about using both MAX and TAF because of changes in data quality and completeness between the two sources. Second, more enrollees would have lookback data in 2021 than 2014 because Medicaid stopped doing redeterminations during the PHE period; this would increase the prevalence of conditions among 2021 enrollees compared with 2014 enrollees because there would be more claims data available for 2021 enrollees, even absent any differences in health status between the populations in those two years. Finally, many enrollees in Medicaid, including newborns or enrollees who gain coverage for pregnancy, will not have lookback data in a prior calendar year.

As a result of using only one year of claims data to construct the CCW-like condition variables, the prevalence of many of these conditions are likely lower for the Medicaid and CHIP population in these analyses than reported in other studies. The condition categories include the following original conditions: acquired hypothyroidism; acute myocardial infarction; Alzheimer's disease and related disorders or senile dementia; anemia; asthma; atrial fibrillation; benign prostatic hyperplasia; cancer – breast; cancer – colorectal; cancer – endometrial; cancer – lung; cancer – prostate; cataract; chronic kidney disease; chronic obstructive pulmonary disease and bronchiectasis; glaucoma; heart failure; hip/pelvic fracture; hyperlipidemia; hypertension; ischemic heart disease; osteoporosis; rheumatoid arthritis/osteoarthritis; and stroke/transient ischemic attack. We also included the following other chronic and potentially disabling conditions: blindness and visual impairment; cystic fibrosis and other metabolic developmental disorders; epilepsy; fibromyalgia, chronic pain and fatigue; hearing impairment; human immunodeficiency virus and/or acquired immunodeficiency syndrome (HIV/AIDS); intellectual disabilities and related conditions; leukemias and lymphomas; migraine and chronic headache; mobility impairments; muscular dystrophy; and peripheral vascular disease. We used these variables in descriptive analyses only (see Exhibit C.4). Exhibit C.3 defines the covariates used in the Medicaid and CHIP regressions for the trend analyses.

Exhibit C.3. Covariates for Medicaid regression models

Covariate	Definition
Age category	Calculated based on the first day observable in Medicaid data for the year (that is, alive, enrolled in Medicaid or CHIP, not dually enrolled in Medicare and Medicaid, eligible for full benefits, and did not have any other third-party coverage)
Ages 0 to 5	= 1 if age >= 0 & age <= 5 = 0 otherwise
Ages 6 to 12	= 1 if age >= 6 & age <= 12 = 0 otherwise
Ages 13 to 18	= 1 if age >= 13 & age <= 18 = 0 otherwise
Ages 19 to 26 (reference category)	= 1 if age >= 19 & age <= 26 = 0 otherwise
Ages 27 to 34	= 1 if age >= 27 & age <= 34 = 0 otherwise
Ages 35 to 44	= 1 if age >= 35 & age <= 44 = 0 otherwise
Ages 45 to 54	= 1 if age >= 45 & age <= 54 = 0 otherwise
Ages 55 and older	= 1 if age >= 55 = 0 otherwise
Sex	
Male (reference category)	= 1 if male or unknown sex = 0 if female
Female	= 1 if female = 0 if male or unknown sex
Major eligibility category	Based on the first month observable in Medicaid data for the year
Pregnancy	= 1 if eligibility group = {05, 53, 67, or 68}
Child	= 2 if age < 21 and eligibility group = {01, 02, 03, 04, 06, 07, 08, 09, 14, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 54, 55, 56, 61, 62, 63, 64, 65, 66, 70, 71, or missing}
Adult non-expansion (reference category)	= 3 if age >= 21 and age < 65 and eligibility group = {01, 02, 03, 04, 09, 14, 27, 32, 33, 34, 35, 36, 56, 70, 71, or missing}
Persons with disabilities	= 4 if age <65 and eligibility group = {11, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23, 25, 26, 37, 38, 39, 40, 41, 42, 43, 44, 46, 51, 52, 59, 60} or if eligibility group = {21, 24, 45, 47, 48, 49, 50, 69}
Adult expansion	= 5 if age >= 18 and eligibility group = {72, 73, 74, 75}
Aged	= 6 if age >= 65 and eligibility group = {01, 02, 03, 04, 05, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 25, 26, 27, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 51, 52, 53, 56, 59, 60, 71, or missing}
Use of home and community-based services waiver or program	= 1 if the beneficiary's eligibility or claims data indicated enrollment in any of the following waivers, authorities, or programs at any time during the year: 1915(c), 1915(i), 1915(j), 1915(k), Money Follows the Person Demonstration, or PACE[a] = 0 otherwise

Covariate	Definition
Zip code-level race and ethnicity	
White (reference category)	Percentage of residents in the enrollee's zip code that reported White race
Black	Percentage of residents in the enrollee's zip code that reported Black race
Hispanic	Percentage of residents in the enrollee's zip code that reported Hispanic ethnicity
Asian	Percentage of residents in the enrollee's zip code that reported Asian race
Other minorities	Percentage of residents in the enrollee's zip code that reported a race other than White, Black, or Asian
Social Vulnerability Index summary score	Mean value of the CDC Social Vulnerability Index Score, calculated across all zip codes that mapped to the same Census tract from the Medicare beneficiary-year analysis file
Rural residence	= 1 if more than 50% of residents in that zip code are living in rural areas, per Census Urban and Rural classification by ZCTA = 0 if 50% or fewer residents in that zip code are living in rural areas, per Census Urban and Rural classification by ZCTA

Source: Mathematica's analyses of Maryland TAF data.

[a] PACE is a program for people dually eligible for Medicare and Medicaid. Because we excluded dually eligible months from the Medicaid analyses, we did not find any enrolled months in PACE among the Medicaid-only population.

CDC = Centers for Disease Control and Prevention; PACE = Program of All-Inclusive Care for the Elderly; TAF = Transformed Medicaid Statistical Information System Analytic File; ZCTA = zip code tabulation area.

C.3. Supplemental analyses

C.3.1. Changes in demographic characteristics from 2014 to 2021

In this section, we report additional descriptive analyses of changes in the composition of the Medicaid population from 2014 to 2021 to show how the population changed over time. The number and percentage of adults enrolled in the adult expansion category increased from 18% in 2014 to 27% in 2021. Relatedly, the percentage of the population in the child category decreased from 54% in 2014 to 49% in 2021. The percentage of enrollees with anxiety disorders more than doubled from 2014 to 2021 from 3% to 7% (Exhibit C.4).

Exhibit C.4. The distribution of demographic and enrollment characteristics and presence of chronic conditions changed from 2014 to 2021 for the Maryland Medicaid and CHIP population, likely driven by growth in the adult expansion population and continuous enrollment requirements during the COVID-19 pandemic

	2014 N = 1,293,428	2021 N = 1,427,675
Demographic and neighborhood characteristics, % unless otherwise noted		
Age category		
0 to 5 years	21	15
6 to 12 years	19	18
13 to 18 years	13	14
19 to 26 years	12	12
27 to 34 years	10	13
35 to 44 years	9	11
45 to 54 years	9	8
55 and older	6	8
Mean age (SD)	21.8 (17.8)	23.9 (17.6)
Female	55	53
Eligibility-related characteristics, %		
Eligibility group		
Child	54	49
Pregnancy	1	1
Adult expansion	18	27
Adult, non-expansion	19	17
Persons with disabilities	8	6
Aged	<1	<1
Enrolled in comprehensive managed care	97	98
Geographic characteristics, %		
Rural	7	7
Geographic characteristics, mean (SD)		
SVI summary score	0.5 (0.2)	0.5 (0.2)
Percentage White	49.8 (28.4)	46.0 (27.6)
Percentage Black	37.7 (28.4)	37.0 (27.1)
Percent Asian	4.9 (5.3)	5.3 (5.9)

	2014 N = 1,293,428	2021 N = 1,427,675
Percentage Hispanic	9.7 (11.1)	11.4 (12.1)
Percentage all other races and ethnicities	7.5 (8.1)	11.7 (10.2)
Presence of chronic conditions, %		
Alzheimer’s, related dementias, or senile dementia	<1	<1
CVD-related conditions		
Acute myocardial infarction	<1	<1
Atrial fibrillation	<1	<1
Heart failure	<1	<1
Hyperlipidemia	3	3
Hypertension	7	6
Ischemic heart disease	1	1
Peripheral vascular disease	<1	<1
Stroke/transient ischemic attack	<1	<1
Conditions associated with higher risk of developing CVD		
Chronic kidney disease	1	2
Diabetes	4	4
Obesity	2	3
Lung-related conditions		
Asthma	5	3
COPD	1	<1
Mental health-related conditions		
ADHD, conduct disorders, and hyperkinetic syndrome	4	4
Anxiety disorders	3	7
Bipolar disorder	3	3
Depressive disorders	8	10
Personality disorders	<1	<1
Post-traumatic stress disorder	<1	2
Schizophrenia and other psychotic disorders	1	1
Substance use-related conditions		
Alcohol use disorders	2	2
Drug use disorders	4	5
Opioid use disorder	4	4
Tobacco use	2	2
Developmental disabilities, intellectual disabilities, learning disabilities, and other chronic and potentially disabling conditions present at birth		
Autism	<1	<1
Cerebral palsy	<1	<1
Cystic fibrosis	<1	<1
Intellectual disabilities	<1	<1
Learning disabilities	3	4
Muscular dystrophy	0	0
Other developmental delays	3	2

	2014 N = 1,293,428	2021 N = 1,427,675
Sickle cell disease	<1	<1
Spina bifida and other congenital anomalies of nervous system	<1	<1
Injuries		
Spinal cord injury	0	<1
Traumatic brain injury and nonpsychotic mental disorders due to brain damage	<1	0
Neurological-related conditions		
Epilepsy	<1	<1
Fibromyalgia, chronic pain, and fatigue	2	2
Migraine and chronic headache	<1	<1
Multiple sclerosis & transverse myelitis	<1	<1
Sensory impairments		
Blindness / visual impairment	0	0
Deafness / hearing impairment	<1	<1
Liver-related conditions		
Liver disease, cirrhosis, and other liver conditions	<1	<1
Viral hepatitis	<1	<1
Musculoskeletal conditions		
Hip/pelvic fracture	0	0
Osteoporosis	<1	<1
Mobility impairments	<1	<1
Rheumatoid arthritis / osteoarthritis	2	2
Eye-related conditions		
Cataract	<1	<1
Glaucoma	<1	1
Other conditions		
Acquired hypothyroidism	1	1
Anemia	5	5
Benign prostatic hyperplasia	<1	<1
Cancer	<1	<1
HIV/AIDS	<1	<1
Pressure and chronic ulcers	<1	<1

Source: Mathematica's analyses of Maryland 2014 and 2021 TAF eligibility and claims data.

Note: We measured the chronic condition variables based on fee-for-service claims and managed care encounter records for each enrollee within each year, 2014 and 2021. The Medicaid and CHIP analysis population excludes people dually enrolled in Medicare and Medicaid (we include these beneficiaries in our Medicare analyses).

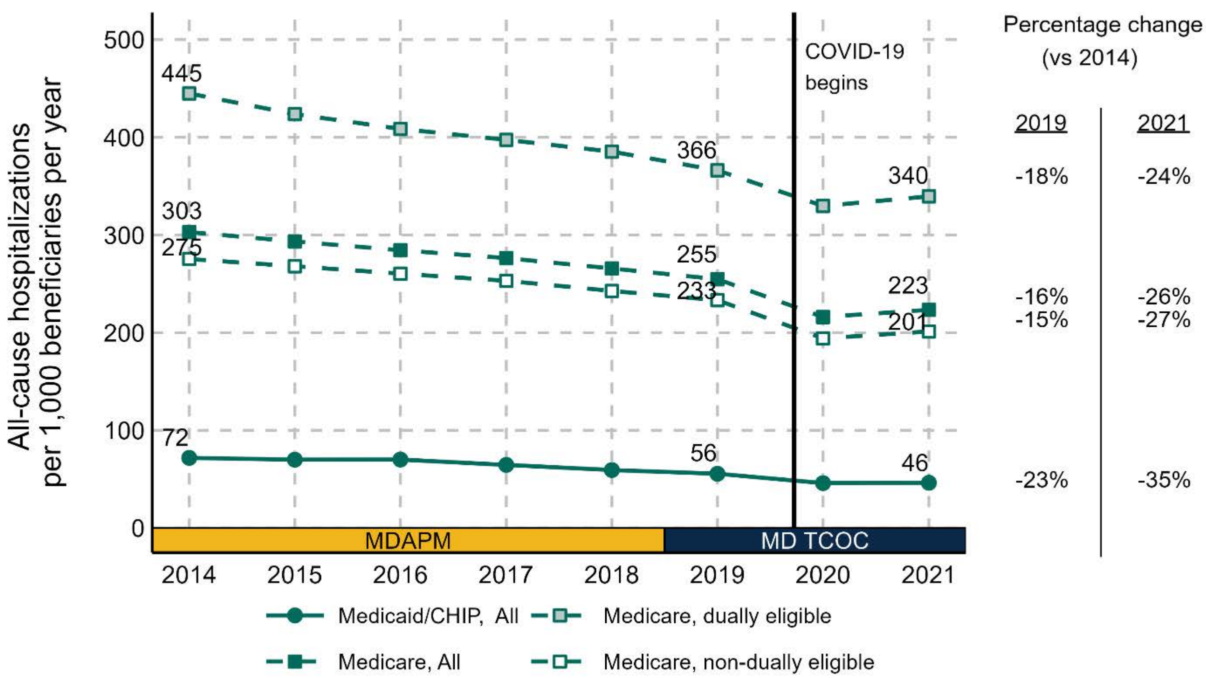
ADHD = attention deficit / hyperactivity disorder; CHIP = Children's Health Insurance Program; COPD = chronic obstructive pulmonary disease; CVD = cardiovascular disease; HIV/AIDS = human immunodeficiency virus/acquired immunodeficiency syndrome; SD = standard deviation; SVI = Social Vulnerability Index; TAF =Transformed Medicaid Statistical Information System Analytic File.

C.3.2. Regression-adjusted trends separated by dual eligible status

We also present regression-adjusted trends for the Maryland Medicare FFS population overall and separately for dually eligible beneficiaries and non-dually eligible beneficiaries to assess whether trends for dually eligible beneficiaries were similar to trends for the Medicaid and CHIP population.

Dually eligible Medicare FFS beneficiaries had higher rates of all-cause hospitalizations in all years as non-dually eligible Medicare FFS beneficiaries but similar percentage reductions over the model period (Exhibit C.5). Medicaid and CHIP enrollees had larger percentage reductions in all-cause hospitalizations, but they had much lower levels of hospitalizations than the Medicare population.

Exhibit C.5. All-cause hospitalizations for Medicaid and CHIP enrollees, all Medicare FFS beneficiaries, and separately for dually and non-dually eligible Medicare FFS beneficiaries

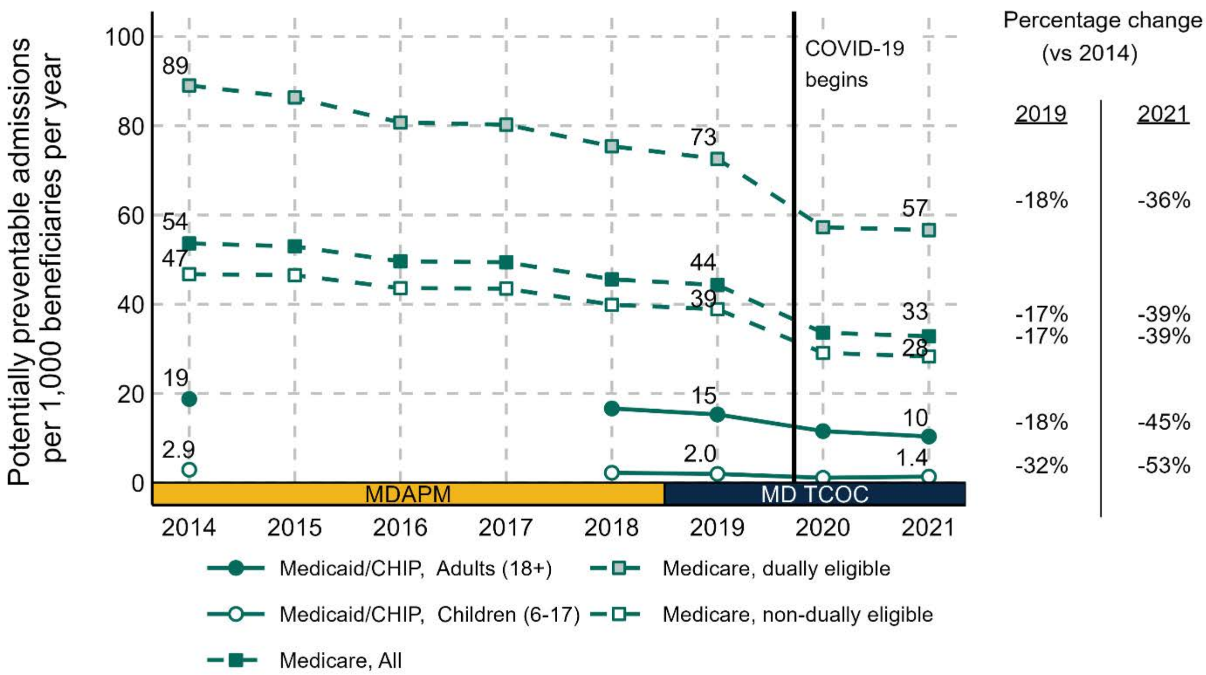


Notes: (1) Mathematica's analyses of Medicare FFS enrollment and inpatient claims and Maryland TAF data; (2) The Medicaid and CHIP analysis population excludes people dually enrolled in Medicare and Medicaid (we include these beneficiaries in the Medicare analyses); (3) All-cause hospitalizations for the Medicaid and CHIP population exclude stays for pregnancy, delivery, or neonates; and (4) In 2021, the analysis populations included N = 1,347,862 Medicaid and CHIP enrollees and N= 737,204 Medicare FFS beneficiaries. The 2021 Medicare FFS population included N=125,314 beneficiaries dually eligible for Medicare and Medicaid and N= 611,891 beneficiaries not dually eligible for Medicare and Medicaid.

CHIP = Children's Health Insurance Program; FFS = fee for service; MDAPM = Maryland All-Payer Model; MD TCOC = Maryland Total Cost of Care; TAF = Transformed Medicaid Statistical Information System Analytic File

Dually eligible Medicare FFS beneficiaries had higher rates of potentially avoidable hospitalizations than non-dually eligible Medicare FFS beneficiaries in all years but similar percentage reductions over the model period. Medicaid and CHIP children ages 6 to 17 had a larger percentage reduction in potentially avoidable hospitalizations, though they had much lower levels than both adult Medicaid and Medicare FFS populations (Exhibit C.6).

Exhibit C.6. Potentially avoidable hospitalizations for Medicaid and CHIP enrollees, all Medicare FFS beneficiaries, and separately for dually and non-dually eligible Medicare FFS beneficiaries

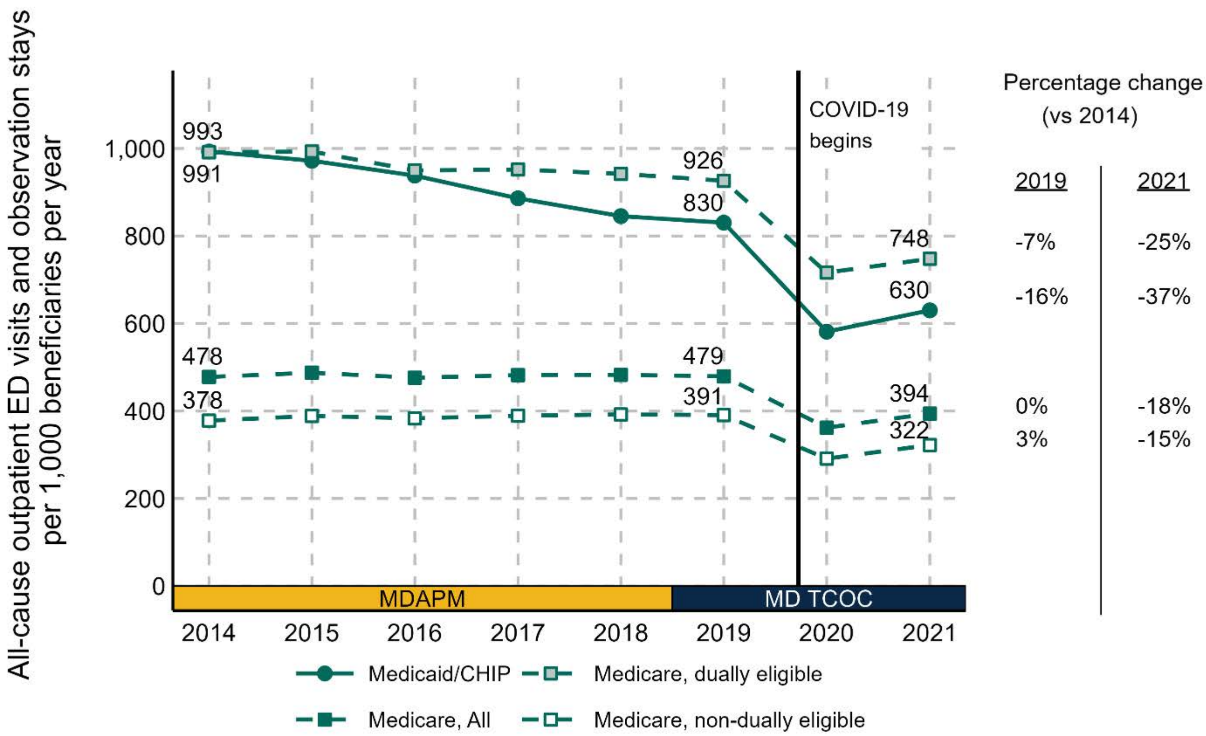


Notes: (1) Mathematica's analyses of Medicare FFS enrollment and inpatient claims and Maryland TAF data; (2) Analyses of potentially avoidable hospitalizations are based on the Agency for Healthcare Research and Quality's PQI and PDI software and are measured for hospitalizations for adults ages 18 and older and for children ages 6 to 17, respectively. For Medicaid PQI and PDI analyses, we did not apply procedure code-based exclusions and we did not construct the PQI for lower extremity amputation among patients with diabetes because of poor quality procedure code data on the Maryland TAF inpatient file from 2014 to 2019; (3) The Medicaid and CHIP analysis population excludes people dually enrolled in Medicare and Medicaid (we include these beneficiaries in the Medicare analyses); (4) In 2021, the analysis populations included N =725,176 Medicaid and CHIP enrollees ages 18 years and older, N = 407,266 Medicaid and CHIP enrollees ages 6 to 17 years, and N= 737,204 Medicare FFS beneficiaries. The 2021 Medicare FFS population included N=125,314 beneficiaries dually eligible for Medicare and Medicaid and N= 611,891 beneficiaries not dually eligible for Medicare and Medicaid.

CHIP = Children's Health Insurance Program; FFS = fee for service; MDAPM = Maryland All-Payer Model; MD TCOC = Maryland Total Cost of Care; PDI = Pediatric Quality Indicator; PQI = Prevention Quality Indicator; TAF = Transformed Medicaid Statistical Information System Analytic File.

Dually eligible Medicare FFS beneficiaries and Medicaid and CHIP enrollees (not dually eligible) had similarly high rates of outpatient ED visits and observation stays at the start of the Maryland Model, but rates declined faster for Medicaid and CHIP enrollees over the model period. Compared with dually eligible beneficiaries and Medicaid and CHIP enrollees, non-dually eligible Medicare FFS beneficiaries had lower levels of outpatient ED use in all years, no change in use from 2014 to 2019, and a smaller dip in 2020 and 2021 (Exhibit C.7).

Exhibit C.7. Outpatient ED visits and observation stays for Medicaid and CHIP enrollees, all Medicare FFS beneficiaries, and separately for dually and non-dually eligible Medicare FFS beneficiaries

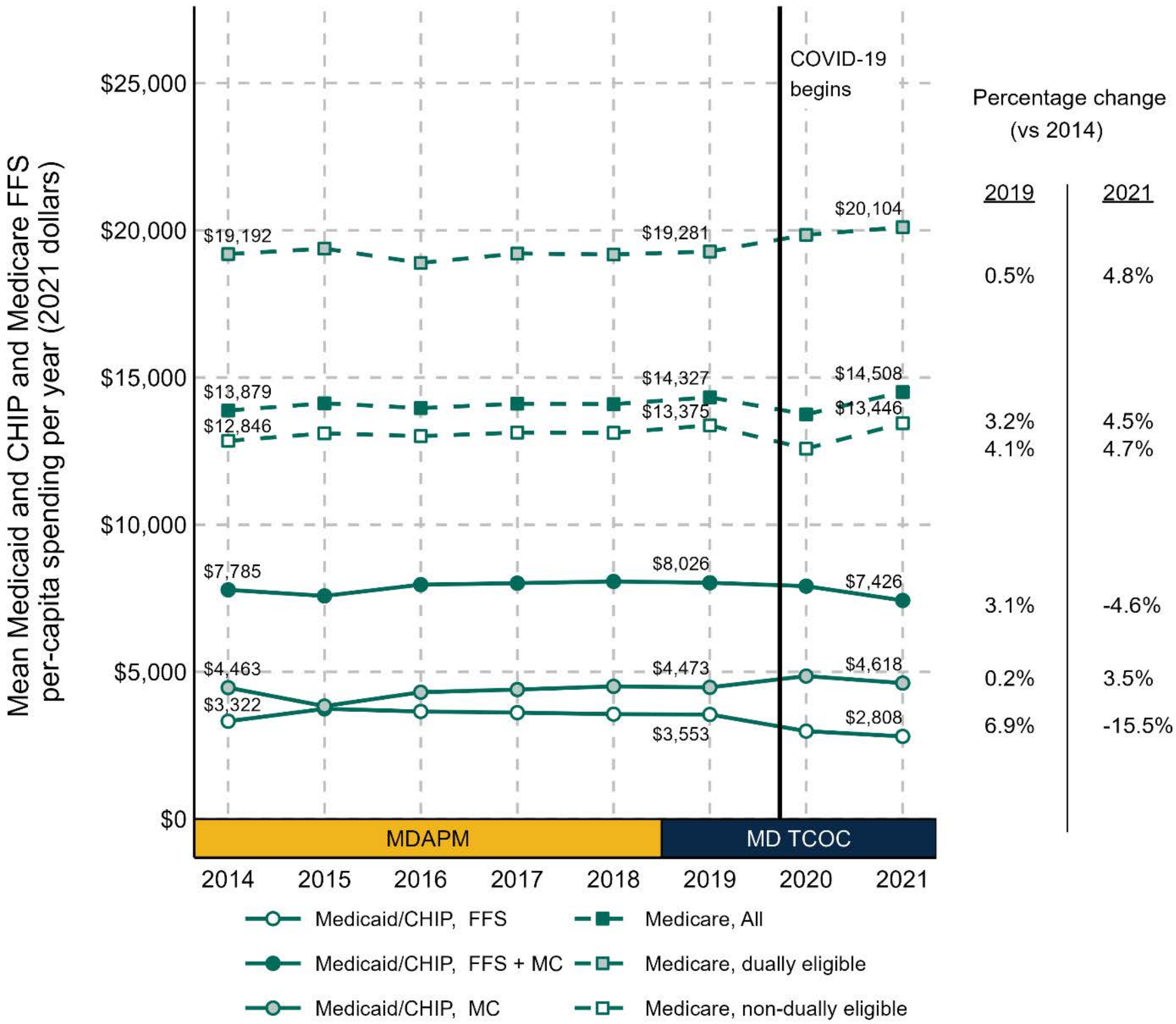


Notes: (1) Mathematica's analyses of Medicare FFS enrollment and inpatient claims and Maryland TAF data; (2) The Medicaid and CHIP analysis population excludes people dually enrolled in Medicare and Medicaid (we include these beneficiaries in the Medicare analyses); (3) In 2021, the analysis populations included N = 1,347,862 Medicaid and CHIP enrollees and N= 737,204 Medicare FFS beneficiaries. The 2021 Medicare FFS population included N=125,314 beneficiaries dually eligible for Medicare and Medicaid and N= 611,891 beneficiaries not dually eligible for Medicare and Medicaid.

CHIP = Children's Health Insurance Program; ED = emergency department; FFS = fee for service; MDAPM = Maryland All-Payer Model; MD TCOC = Maryland Total Cost of Care; TAF = Transformed Medicaid Statistical Information System Analytic File.

Dually eligible Medicare FFS beneficiaries had higher per-capita spending in all years and lower growth in spending from 2014 to 2019 than non-dually eligible Medicare FFS beneficiaries and Medicaid and CHIP enrollees. From 2014 to 2021, however, growth in per-capita spending was the same for dually eligible and non-dually eligible Medicare FFS beneficiaries (Exhibit C.8).

Exhibit C.8. Per-capita spending for Medicaid and CHIP enrollees, all Medicare FFS beneficiaries, and separately for dually and non-dually eligible Medicare FFS beneficiaries



Notes: (1) Mathematica's analyses of Medicare FFS enrollment and inpatient claims and Maryland TAF data; (2) We used the Gross Domestic Product price index to inflate spending from 2014 to 2020 into 2021 dollars. For more information, see Using Appropriate Price Indices for Expenditure Comparisons (ahrq.gov); (3) The Medicaid and CHIP analysis population excludes people dually enrolled in Medicare and Medicaid (we include these beneficiaries in the Medicare analyses); (4) In 2021, the analysis populations included N = 1,347,862 Medicaid and CHIP enrollees and N= 737,204 Medicare FFS beneficiaries. The 2021 Medicare FFS population included N=125,314 beneficiaries dually eligible for Medicare and Medicaid and N= 611,891 beneficiaries not dually eligible for Medicare and Medicaid.

CHIP = Children's Health Insurance Program; ED = emergency department; FFS = fee for service; MC = managed care; MDAPM = Maryland All-Payer Model; MD TCOC = Maryland Total Cost of Care; TAF = Transformed Medicaid Statistical Information System Analytic File.

Appendix D. Methods and Supplemental Results for Effects of the Model on Health Equity for Medicare Fee-For-Service Beneficiaries

In Chapter 4 of the Progress Report, we estimate whether the Maryland Model improved health equity by reducing disparities (or gaps) in quality-of-care measures by race (comparing Black and White beneficiaries) and place (based on the Centers for Disease Control and Prevention's Social Vulnerability Index (SVI). This appendix provides more details about the subgroups and outcomes used in Chapter 4 (Section D.1), baseline assessments of outcome levels and covariate balance (Section D.2), detailed descriptions of the regression approach (Section D.3), and supplemental results (Section D.4).

D.1. Subgroups and outcomes

D.1.1. Rationale for choosing subgroups and outcomes

In consultation with CMS, we chose two subgroups and three quality-of-care outcomes (a total of six subgroup-outcome combinations), described in Chapter 4. Each of these choices aligned with the model logic either because (a) the model contained specific goals, incentives, and/or supports to reduce disparities now or in the future or (b) the model had the potential to impact the outcome and we wanted to assess whether those impacts were experienced equally across subgroups. We first identified a limited set of high-priority subgroups and outcomes in 2021, prior to selecting a comparison group for the statewide analysis. Pre-specifying the subgroups and outcomes enabled us to achieve better balance on a limited set of factors. Narrowing the number of subgroup-outcome combinations also minimized the risk of finding chance differences due simply to multiple testing. After selecting a comparison group, we added timely follow-up as a key outcome based on work that the Health Services Cost Review Commission (HSCRC) did in 2022 to identify disparities in timely follow-up in Maryland (HSCRC 2022b). Exhibit D.1 describes our rationale for choosing each of the three quality-of-care measures.

Exhibit D.1. Rationale for choosing quality-of-care measures

Quality-of-care measure	Rationale
Potentially preventable admissions	A decrease in the rates of potentially preventable admissions is one of the quality goals that CMS and Maryland set in SIHIS (HSCRC 2020a). Reducing potentially preventable admissions could signal improvements in patient health status resulting from improved ambulatory care, better care coordination, and other factors. We wanted to assess whether any reductions in potentially preventable admissions were experienced equally by race and place. Previous studies have also shown large disparities in potentially preventable admissions by race and region (Mukamel et al. 2015; Billings et al. 1996).
30-day unplanned readmissions	CMS and Maryland set a goal in SIHIS to reduce disparities in readmissions (HSCRC 2020a), which materialized into incentive payments to hospitals for reducing within-hospital disparities in 30-day unplanned readmissions in 2022 after a delay due to the COVID-19 pandemic. Specifically, the program assesses patient-level socioeconomic exposure using the Patient Adversity Index, which varies by combinations of race, Medicaid status, and ADI national percentiles, and rewards hospitals reducing socioeconomic disparities in readmission by providing additional payments (up to 0.5% of inpatient revenue; HSCRC n.d.). Given these recent incentives, we would expect to see differential impacts by race and place moving forward.
Timely follow-up after acute exacerbation of a chronic condition	In 2022, the HSCRC described timely follow-up rates stratified by race, dual eligibility status, and ADI. It noted disparities on all three factors within Maryland—that is, beneficiaries who are Black, dually eligible, and those living in areas with higher deprivation had higher odds of not receiving follow-up care compared to patients without such attributes. Given these results, the HSCRC plans to develop hospital incentives for reducing disparities on this measure, similar to the approved policy on improving readmission disparity gaps (HSCRC 2022b). Thus, as in the case of the readmission outcome, we would expect to see differential impacts across subgroups in the timely follow-up measure moving forward.

ADI = Area Deprivation Index; HSCRC = Health Services Cost Review Commission; SIHIS = Statewide Integrated Health Improvement Strategy.

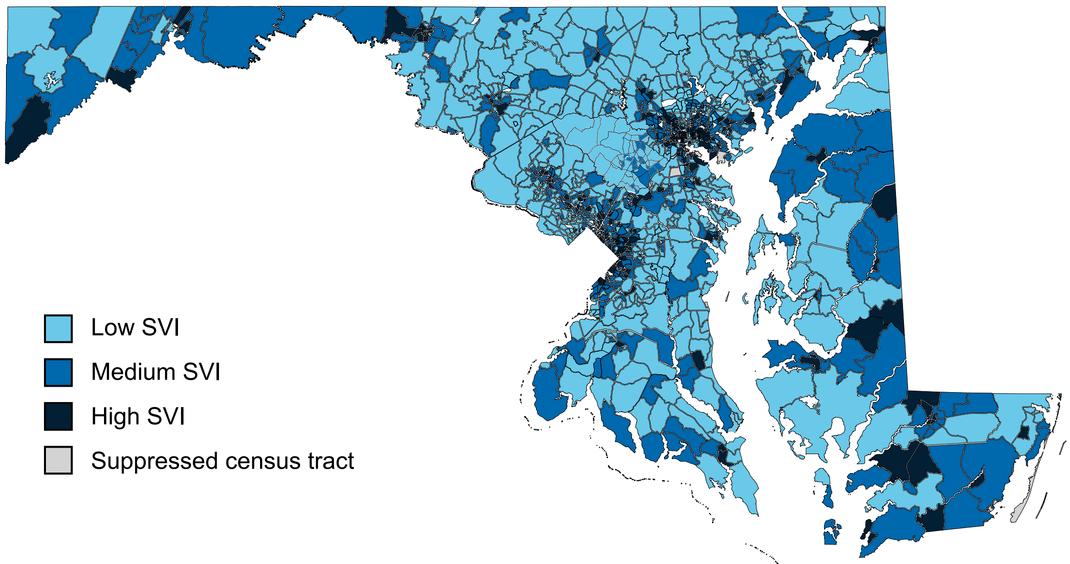
We chose to examine differences by race and place to align with the goals of CMS and Maryland's Statewide Integrated Health Improvement Strategy (SIHIS). We chose SVI as a measure of place due to the following advantages:

1. SVI is a validated, consistent, and accepted measure of vulnerability widely and easily available in all years of our analysis, going back more than 10 years. It performs as well as other indices of neighborhood deprivation, including the Area Deprivation Index (ADI), at identifying disparities in health care (Carmichael et al. 2020). The SVI and ADI, as well as other similar measures of neighborhood deprivation, use the Census American Community Survey (ACS) as the primary data source and tend to rank neighborhoods similarly, though there are differences (Exhibit D.2).
2. The SVI was used in matching and as a regression control in the statewide impacts analysis, allowing for consistency between the statewide impacts analysis and subgroup analyses.
3. Using the SVI, which is composed of four domains including socioeconomic status and transportation, will enable us in the future to conduct domain-specific analyses for a finer look at where impacts are strongest.

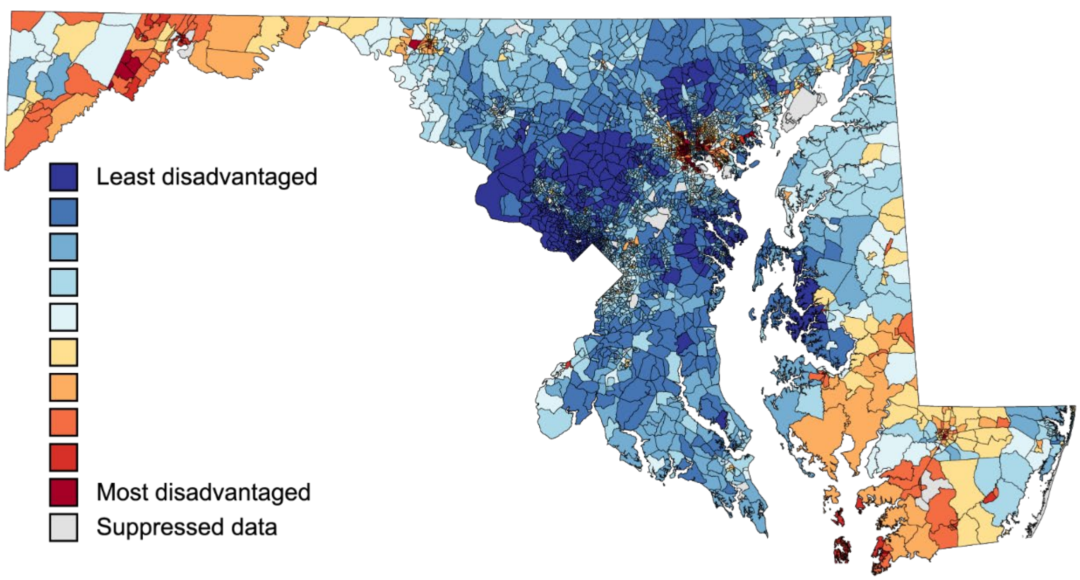
4. Recent critiques of the ADI, one of the alternative measures, suggest that ADI is heavily weighted towards median home values and thus might represent less of a multidimensional vulnerability score than SVI (Hannan et al. 2023).

Exhibit D.2. Comparison of the Social Vulnerability Index and Area Deprivation Index

Panel A: Social Vulnerability Index



Panel B: Area Deprivation Index



Source: Panel A is 2018 CDC/ATSDR SVI mapped to 2010 census tracts. Panel B is 2021 national ADI sourced from the University of Madison-Madison website: https://www.neighborhoodatlas.medicine.wisc.edu/.

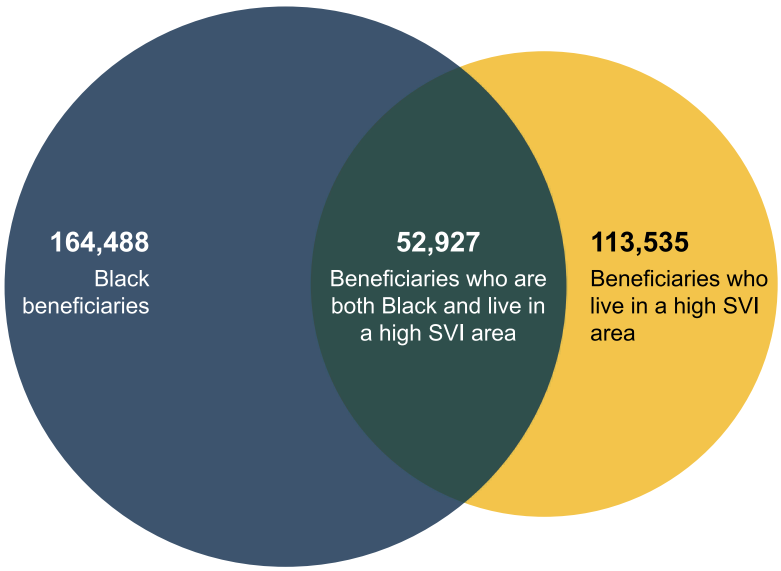
ADI = Area Deprivation Index; ATSDR = Agency for Toxic Substances and Disease Registry; CDC = Centers for Disease Control and Prevention; SVI = Social Vulnerability Index

Prior to finalizing the selection of subgroup-outcome combinations, we assessed baseline balance and trends (see Section D.2 for more details) and confirmed sufficient power to detect subgroup differences.

D.1.2. Overlap of race and place

There is some, but not complete, overlap between race and place (Exhibit D.3). For example, in 2022 in Maryland, Black beneficiaries comprised 47% of beneficiaries living in high SVI areas but only 13% of beneficiaries living in low SVI areas. Similarly, 32% of Black beneficiaries lived in high SVI areas compared to only 10% of White beneficiaries.

Exhibit D.3. Overlap between Black race and residence in high SVI areas, Maryland 2022



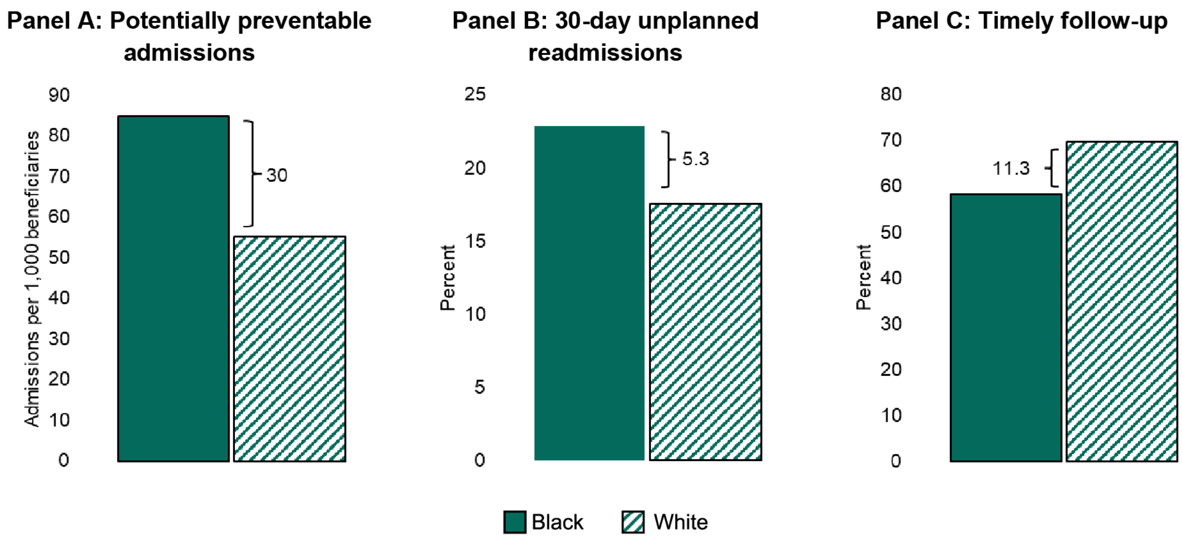
SVI = Social Vulnerability Index.

D.2. Baseline assessments

D.2.1. Baseline differences in outcomes

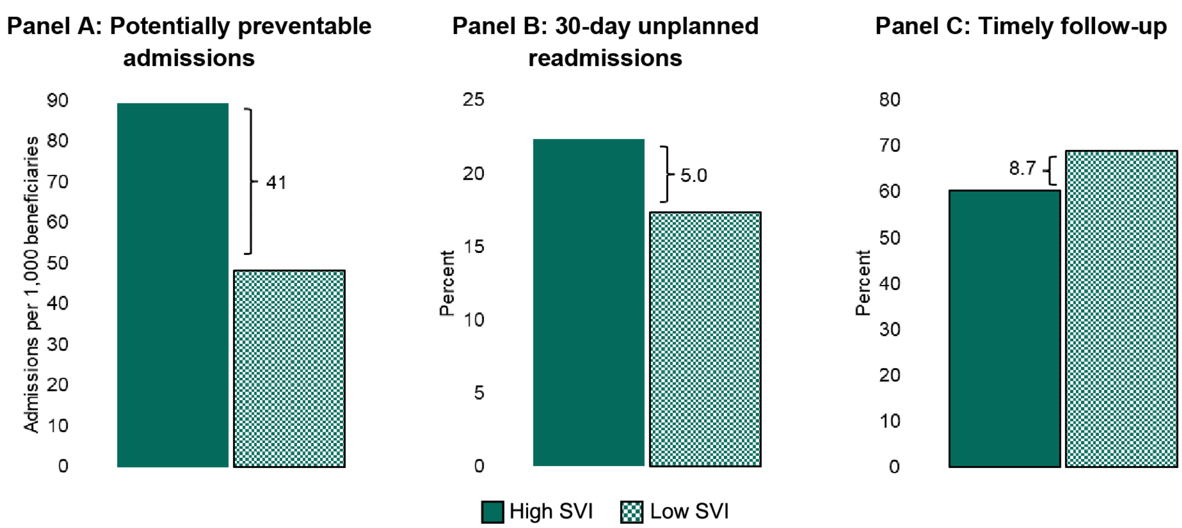
There were substantial disparities in the three quality-of-care measures at baseline (2011–2013). Black beneficiaries, relative to White beneficiaries, had 30 more potentially preventable admissions per 1,000 beneficiaries (54% higher), 5.3 percentage points (30%) higher readmission rates, and 11.3 percentage points (16%) fewer timely follow-ups (Exhibit D.4). Beneficiaries living in high, relative to low, SVI areas had 41 more potentially preventable admissions per 1,000 beneficiaries (85% higher), 5.0 percentage points (29%) higher readmission rates, and 8.7 percentage points (13%) fewer timely follow-ups (Exhibit D.5).

Exhibit D.4. Lower quality of care for Maryland Medicare FFS Black beneficiaries at baseline (2011–2013)



FFS = fee for service.

Exhibit D.5. Lower quality of care for Maryland Medicare FFS beneficiaries living in high vulnerability areas at baseline (2011–2013)



FFS = fee for service; SVI = Social Vulnerability Index.

D.2.2. Balance within and across subgroups

We assessed baseline balance between Maryland and comparison regions on the subgroups of interest, as well as balance between Maryland and comparison regions on key covariates within the subgroups (Exhibit D.6). Maryland and comparison regions had a similar proportion of beneficiaries in each race subgroup, which was expected because this subgroup was tightly matched on. However, Maryland had a smaller percentage of beneficiaries living in more socially vulnerable areas than other states. Within subgroups, we achieved good balance on baseline characteristics for White beneficiaries but observed relatively large differences in key beneficiary characteristics between Maryland and the comparison group among Black beneficiaries. We believe that regression adjustment mitigated these differences.

Exhibit D.6. Baseline balance on subgroups and high-priority characteristics within each subgroup (2013)

Variable description	Maryland post-weighted mean	Selected comparison group (national post-weighted mean)	Standardized difference post-weighting
Black and White beneficiaries (%)			
Black	24.37	22.37	0.08
White	75.63	77.63	-0.08
High-priority characteristics among Black beneficiaries[a]			
Average age	67.65	66.59	0.40
Percentage female	59.17	58.41	0.21
Percentage with original reason for Medicare entitlement: disability, ESRD	35.20	42.36	-0.74
Average HCC risk score	1.34	1.39	-0.33
High-priority characteristics among White beneficiaries[a]			
Average age	72.78	72.77	0.01
Percentage female	57.07	56.75	0.14
Percentage with original reason for Medicare entitlement: disability, ESRD	15.66	16.48	-0.10
Average HCC risk score	1.05	1.06	-0.04
Beneficiaries living in high and low SVI areas (%)			
High SVI	17.32	26.38	-0.35
Low SVI	50.50	43.09	0.26
High-priority characteristics among beneficiaries living in high SVI areas[a]			
Average age	68.63	68.56	0.03
Percentage female	58.47	57.60	0.31
Percentage with original reason for Medicare entitlement: disability, ESRD	34.76	35.86	-0.13
Average HCC risk score	1.32	1.30	0.12

Variable description	Maryland post-weighted mean	Selected comparison group (national post-weighted mean)	Standardized difference post-weighting
High-priority characteristics among beneficiaries living in low SVI areas[a]			
Average age	72.42	72.73	-0.20
Percentage female	56.51	56.14	0.16
Percentage with original reason for Medicare entitlement: disability, ESRD	13.90	13.70	0.03
Average HCC risk score	1.02	1.02	0.02

Note: Post-weighted means are weighted by the final matching weights. Standardized differences are the difference between treatment and comparison group means, divided by the PUMA-level pooled standard deviation. High SVI is defined as the highest (most vulnerable) tertile based on national levels, and low SVI is defined as the lowest (least vulnerable) tertile.

[a] These beneficiary characteristics are not subgroups or outcomes, but we required 0.25 standardized differences during matching for the overall sample, given their importance. We do not expect close similarity on all these characteristics (as we did not match on them directly within each subgroup, except the average HCC score for Black and White beneficiaries) but will use this information to interpret findings.

ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; PUMA = Public Use Microdata Area; SVI = Social Vulnerability Index.

D.2.3. Balance on outcome levels and trends

We assessed balance on levels and trends in the three quality-of-care measures within subgroups to identify any deviations from the assumption of parallel outcome trends which could lead to bias in our difference-in-differences models (see details about the regression approach in section D.3). There was overall acceptable balance on trends for potentially preventable admissions and 30-day unplanned readmissions, with none exceeding the 0.33 Public Use Microdata Area (PUMA)-level standardized differences on the 2011−2013 trends, the tolerance level we set for some of the other key subgroup variables in matching. However, there was somewhat poorer balance on timely follow-up after acute exacerbation of chronic conditions, particularly within SVI subgroups (Exhibit D.7). We conducted additional regression-based tests for possible violations of the assumption of our difference-in-differences regression models and found it was unlikely that imbalances would lead to a violation in the assumption of parallel trends (results not shown).

Exhibit D.7. Baseline balance on levels and trends in quality-of-care measures

Variable description	Maryland post-weighted mean	Selected comparison group (national post-weighted mean)	Standardized difference post-weighting
Among Black beneficiaries			
Potentially preventable admissions, 2013	80.03	76.82	0.17
Potentially preventable admissions, 2011–2013 trend[a]	-4.64	-5.49	0.11
30-day post-discharge unplanned readmission, 2013	21.87	21.97	-0.03
30-day post-discharge unplanned readmission, 2011–2013 trend[a]	-0.89	-0.67	-0.13
Timely follow-up after acute exacerbation of chronic conditions, 2013	60.09	58.34	0.29
Timely follow-up after acute exacerbation of chronic conditions, 2011–2013 trend[a]	1.69	0.91	0.28
Among White beneficiaries			
Potentially preventable admissions, 2013	51.89	51.63	0.02
Potentially preventable admissions, 2011–2013 trend[a]	-4.26	-3.84	-0.14
30-day post-discharge unplanned readmission, 2013	16.70	15.97	0.36
30-day post-discharge unplanned readmission, 2011–2013 trend[a]	-0.80	-0.66	-0.21
Timely follow-up after acute exacerbation of chronic conditions, 2013	70.57	71.00	-0.09
Timely follow-up after acute exacerbation of chronic conditions, 2011–2013 trend[a]	0.70	1.11	-0.29
Among beneficiaries living in high SVI areas			
Potentially preventable admissions, 2013	83.31	75.28	0.45
Potentially preventable admissions, 2011–2013 trend[a]	-5.72	-5.04	-0.15
30-day post-discharge unplanned readmission, 2013	20.85	20.22	0.22
30-day post-discharge unplanned readmission, 2011–2013 trend[a]	-1.07	-0.81	-0.26
Timely follow-up after acute exacerbation of chronic conditions, 2013	62.82	61.52	0.22
Timely follow-up after acute exacerbation of chronic conditions, 2011–2013 trend[a]	1.55	0.88	0.37

Variable description	Maryland post-weighted mean	Selected comparison group (national post-weighted mean)	Standardized difference post-weighting
Among beneficiaries living in low SVI areas			
Potentially preventable admissions, 2013	45.07	43.03	0.17
Potentially preventable admissions, 2011–2013 trend[a]	-3.75	-3.39	-0.11
30-day post-discharge unplanned readmission, 2013	16.59	15.53	0.49
30-day post-discharge unplanned readmission, 2011–2013 trend[a]	-0.69	-0.63	-0.06
Timely follow-up after acute exacerbation of chronic conditions, 2013	69.35	71.35	-0.40
Timely follow-up after acute exacerbation of chronic conditions, 2011–2013 trend[a]	0.33	1.08	-0.38

Notes: Post-weighted means are weighted by the final matching weights. Standardized differences are the difference between treatment and comparison group means, divided by the PUMA-level pooled standard deviation.

[a] We used the following process to calculate subgroup-specific PUMA-level trends in outcomes from 2011 to 2013: (1) For each PUMA-subgroup combination, we calculated the level of the outcome measure in the years 2011, 2012, and 2013. (2) We estimated a simple ordinary least squares regression for each PUMA for each outcome measure separately by subgroup. Each regression had just three observations and two explanatory variables—the year (1, 2, or 3) and a constant term. (3) We recorded the coefficient on the year variable for each PUMA-subgroup combination for each outcome measure. This gave us a data set with one row per PUMA per subgroup and columns containing the slope term from each regression (one column per outcome measure). (4) For each outcome measure-subgroup combination, we then calculated the average of these slope measures across (a) the PUMAs in Maryland and (b) the PUMAs in other states and the differences in means, weighted by their subgroup PUMA sample size relevant to that measure (for example, discharges for 30-day unplanned readmissions, all FFS beneficiaries for preventable admissions).

FFS = fee for service; PUMA = Public Use Microdata Area; SVI = Social Vulnerability Index.

D.3. Regression approach

D.3.1. Regression model

We used the same general difference-in-differences model regression approach as described in Chapter 2 and Appendix B but included additional triple interaction terms between (a) indicators for the subgroup of interest (Black or high SVI), (b) treatment status, and (c) each intervention year. This approach allowed estimation of impacts, relative to baseline (2011−2013), for each subgroup (for example, impacts for Black beneficiaries) as well as differences in impacts between subgroups (for example, differences between Black and White beneficiaries). We also included selected interactions between the subgroups and other important covariates, such as age, enrollment time, and some chronic conditions. As in Chapter 2, we estimated impacts on potentially preventable admissions and timely follow-up with beneficiary-level models and 30-day unplanned readmissions with discharge-level models, each including PUMA-level fixed effects.

D.3.2. Estimating effects on quality gaps

Estimating effects on the quality gaps between subgroups (Black–White and high–low SVI) helps to interpret the impact estimates on a more policy-relevant scale. We calculated the effects on gaps using the following steps:

1. We calculated the observed (unadjusted) means for the quality measure in each subgroup during the early Maryland All-Payer Model (MDAPM) period (2014−2016), the late MDAPM period (2017−2018), and the MD TCOC period (2019−2022). For example, during the MD TCOC period, Black beneficiaries in Maryland had, on average, 52 potentially preventable admissions per 1,000 beneficiaries compared to 32 for White beneficiaries— a gap of 20 admissions.
2. We estimated what these means would have been in the absence of the model (the counterfactual) by taking the observed means and subtracting the impact estimates for each subgroup. For example, the impact estimates during MD TCOC, relative to the baseline period, were -17 potentially preventable admissions per 1,000 Black beneficiaries and -4 potentially preventable admissions per 1,000 White beneficiaries (Exhibit D.8). So, we estimate that the means would have been 69 admissions per 1,000 (52 – (-17)) for Black beneficiaries and 36 admissions per 1,000 (32 – (-4)) for White beneficiaries in the absence of the model— a gap of 33 admissions.
3. We compared the observed difference in means by subgroup (the actual gap) and counterfactual differences by subgroup (the expected gap without the model). For example, the model narrowed the gap in potentially preventable admissions from 33 admissions to 20 admissions— an absolute change of 13 admissions and a relative change of 40% (13 / 33, after rounding).

D.4. Supplemental results

D.4.1. Detailed impact estimates

In this section, we present impact estimates of the Maryland Model within and across subgroups by year. Using impacts on potentially preventable admissions by race as an example (Exhibit D.8), the following is a description of how readers can interpret the exhibits in this section:

- Among Black beneficiaries, potentially preventable admissions declined faster in Maryland than for the comparison between baseline (2011−2013) and the first year of MD TCOC (2019), leading to a difference-in-differences estimate of -17.9 per 1,000 beneficiaries per year. The 90% confidence interval around this estimate is -21.2 to -14.7, reflecting that the estimate is statistically different from zero ($p < 0.05$). This is a 22.2% reduction compared to what we would expect in absence of the model.[37] We calculated the impacts in 2020, 2021, and 2022 the same way.

[37] The percentage equals the impact estimate divided by the estimated counterfactual (which equals the Maryland mean minus the impact estimate).

- Combining the four estimates from 2019, 2020, 2021, and 2022, we reach a similar effect during the four years of the MD TCOC period of -17.2 potentially preventable admissions per 1,000 beneficiaries, which is statistically significant.
- We calculated the difference in estimates during the MD TCOC period and later MDAPM period in the same way, but we used the combined later MDAPM period estimates as the baseline. This allowed us to estimate the additional impact that occurred during just the MD TCOC period. Using 2022 as an example, a difference-in-differences estimate of -15.5 admissions per 1,000 during MD TCOC relative to the -13.2 admissions per 1,000 reduction during later MDAPM led to a difference in estimates of -2.4 admissions per 1,000 beneficiaries (-15.5 – (-13.2)), which is statistically significant ($p < 0.05$) with a 90% confidence interval (CI) of -4.3 to -0.4.
- We calculated the same quantities for White beneficiaries. We also calculated triple-difference estimates comparing estimates for White and Black beneficiaries. Using 2022 as an example, we compared the difference-in-differences estimate of -15.5 admissions per 1,000 for Black beneficiaries to the estimate of -4.1 admissions per 1,000 for White beneficiaries, for a triple-difference estimate of -11.4 admissions per 1,000, which is statistically significant ($p < 0.05$) with a 90% CI of -14.4 to -8.4. In other words, the model reduced potentially preventable admissions more for Black than for White beneficiaries, by 11.4 admissions per 1,000.

Exhibit D.8. Impacts of the Maryland Model on quality-of-care outcomes by race

	Difference-in-differences impact estimate for Black beneficiaries		Difference-in-differences impact estimate for White beneficiaries		Triple-difference impact estimate for Black – White beneficiaries	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
Potentially preventable admissions (number per 1,000 beneficiaries per year)						
Early MDAPM period						
2014	-3.3* (-6.1; -0.5)	-4.3%	0.7 (-0.5; 2.0)	1.5%	-4.0** (-6.9; -1.1)	-5.8 pp
2015	-3.2* (-6.1; -0.3)	-4.2%	0.7 (-0.8; 2.2)	1.5%	-3.9** (-6.8; -1.0)	-5.7 pp
2016	-5.8** (-8.6; -2.9)	-7.9%	0.8 (-0.8; 2.4)	1.8%	-6.6** (-9.5; -3.6)	-9.7 pp
Later MDAPM period						
2017	-11.9** (-14.9; -8.9)	-15.0%	-1.0 (-2.8; 0.8)	-2.2%	-10.9** (-13.6; -8.1)	-12.8 pp
2018	-14.4** (-17.2; -11.6)	-18.2%	-4.1** (-7.0; -1.3)	-9.2%	-10.3** (-13.1; -7.4)	-9.0 pp
Combined (2017–2018)	-13.2** (-15.8; -10.5)	-16.6%	-2.6* (-4.8; -0.3)	-5.7%	-10.6** (-13.1; -8.1)	-10.9 pp
MD TCOC period						
2019	-17.9** (-21.2; -14.7)	-22.2%	-4.7** (-7.3; -2.1)	-10.6%	-13.2** (-16.5; -9.9)	-11.6 pp
2020	-17.9** (-20.8; -15.0)	-27.2%	-3.8** (-6.3; -1.3)	-11.4%	-14.1** (-17.1; -11.2)	-15.8 pp
2021	-17.2** (-20.6; -13.8)	-26.7%	-4.3** (-7.1; -1.5)	-13.1%	-12.9** (-16.1; -9.7)	-13.6 pp
2022	-15.5** (-18.5; -12.6)	-24.8%	-4.1** (-7.0; -1.2)	-12.3%	-11.4** (-14.4; -8.4)	-12.5 pp
Combined (2019–2022)	-17.2** (-20.1; -14.3)	-25.1%	-4.2** (-6.8; -1.6)	-11.7%	-13.0** (-15.8; -10.1)	-13.4 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	-4.8** (-6.8; -2.7)	-5.6 pp	-2.1** (-3.0; -1.2)	-4.9 pp	-2.6** (-4.7; -0.5)	-0.7 pp
2020	-4.7** (-7.0; -2.5)	-10.6 pp	-1.2 (-2.5; 0.1)	-5.7 pp	-3.6** (-5.9; -1.3)	-4.9 pp
2021	-4** (-6.6; -1.5)	-10.1 pp	-1.7** (-2.9; -0.5)	-7.4 pp	-2.3* (-4.7; 0)	-2.7 pp
2022	-2.4** (-4.3; -0.4)	-8.2 pp	-1.5* (-3.1; 0)	-6.6 pp	-0.8 (-2.8; 1.2)	-1.6 pp
Combined (2019–2022)	-4** (-5.9; -2.2)	-8.5 pp	-1.6** (-2.7; -0.6)	-6.0 pp	-2.4** (-4.1; -0.6)	-2.5 pp

	Difference-in-differences impact estimate for Black beneficiaries		Difference-in-differences impact estimate for White beneficiaries		Triple-difference impact estimate for Black – White beneficiaries	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
30-day post-discharge unplanned readmissions (percentage of discharges)						
Early MDAPM period						
2014	-0.7** (-1.1; -0.2)	-3.2%	-0.1 (-0.3; 0.2)	-0.6%	-0.6** (-1.1; -0.1)	-2.6 pp
2015	-1.1** (-1.6; -0.6)	-5.1%	-0.5** (-0.8; -0.2)	-3.1%	-0.6* (-1.1; -0.1)	-2.0 pp
2016	-1.3** (-1.8; -0.7)	-6.1%	-1.0** (-1.2; -0.8)	-6.2%	-0.3 (-0.8; 0.3)	0.1 pp
Later MDAPM period						
2017	-2.5** (-3.2; -1.9)	-11.5%	-1.0** (-1.3; -0.7)	-6.2%	-1.5** (-2.2; -0.8)	-5.3 pp
2018	-2.3** (-2.9; -1.6)	-10.6%	-1.0** (-1.3; -0.7)	-6.2%	-1.2** (-1.9; -0.6)	-4.4 pp
Combined (2017–2018)	-2.4** (-3.0; -1.8)	-11.1%	-1.0** (-1.2; -0.8)	-6.2%	-1.4** (-2.0; -0.8)	-4.9 pp
MD TCOC period						
2019	-3.3** (-4.0; -2.7)	-15.1%	-1.3** (-1.6; -0.9)	-8.0%	-2.1** (-2.8; -1.4)	-7.1 pp
2020	-3.2** (-3.8; -2.6)	-14.7%	-1.3** (-1.6; -0.9)	-7.9%	-2** (-2.6; -1.3)	-6.8 pp
2021	-2.5** (-3.0; -2.0)	-11.5%	-1.0** (-1.3; -0.7)	-6.1%	-1.6** (-2.1; -1.0)	-5.4 pp
2022	-2.6** (-3.2; -2.0)	-12.0%	-0.9** (-1.3; -0.5)	-5.5%	-1.7** (-2.4; -1.0)	-6.5 pp
Combined (2019–2022)	-3.0** (-3.5; -2.5)	-13.8%	-1.1** (-1.4; -0.9)	-6.7%	-1.9** (-2.3; -1.4)	-7.1 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	-1.0** (-1.4; -0.5)	-4.0 pp	-0.3 (-0.6; 0)	-1.8 pp	-0.7** (-1.2; -0.1)	-2.2 pp
2020	-0.8** (-1.4; -0.3)	-3.6 pp	-0.3 (-0.6; 0.1)	-1.7 pp	-0.6 (-1.2; 0.1)	-1.9 pp
2021	-0.1 (-0.7; 0.5)	-0.4 pp	0 (-0.2; 0.3)	0.1 pp	-0.2 (-0.8; 0.5)	-0.5 pp
2022	-0.2 (-0.8; 0.4)	-0.9 pp	0.1 (-0.2; 0.5)	0.7 pp	-0.3 (-1.0; 0.3)	-1.6 pp
Combined (2019–2022)	-0.6** (-1; -0.1)	-2.7 pp	-0.1 (-0.3; 0.1)	-0.5 pp	-0.5* (-0.9; 0)	-2.2 pp
Timely follow-up after acute exacerbation of chronic conditions (percentage of discharges)						
Early MDAPM period						
2014	0.1 (-0.9; 1.0)	0.2%	-0.6 (-1.1; 0)	-0.8%	0.6 (-0.5; 1.8)	1.0 pp
2015	1.9** (1.0; 2.7)	3.1%	0.7** (0.2; 1.3)	1.0%	1.1** (0.2; 2.1)	2.1 pp
2016	3.0** (2.1; 4.0)	4.9%	0.6 (-0.1; 1.4)	0.8%	2.4** (1.3; 3.5)	4.1 pp
Later MDAPM period						
2017	2.7** (1.8; 3.5)	4.4%	0.8 (0; 1.6)	1.1%	1.9** (0.8; 3.0)	3.3 pp
2018	2.8** (1.8; 3.8)	4.5%	0.3 (-0.4; 1.1)	0.4%	2.4** (1.3; 3.6)	4.1 pp
Combined (2017–2018)	2.7** (1.9; 3.5)	4.3%	0.6 (-0.2; 1.3)	0.8%	2.2** (1.2; 3.2)	3.5 pp

	Difference-in-differences impact estimate for Black beneficiaries		Difference-in-differences impact estimate for White beneficiaries		Triple-difference impact estimate for Black – White beneficiaries	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
MD TCOC period						
2019	3.2** (2.3; 4.1)	5.1%	0.7 (-0.1; 1.5)	0.9%	2.5** (1.5; 3.6)	4.2 pp
2020	3.7** (2.5; 4.9)	6.3%	0.6 (-0.3; 1.4)	0.8%	3.1** (1.9; 4.3)	5.5 pp
2021	4.0** (3.0; 5.0)	6.6%	0.2 (-0.6; 1.0)	0.3%	3.8** (2.6; 5.1)	6.3 pp
2022	4.4** (3.2; 5.6)	7.2%	0.5 (-0.5; 1.5)	0.7%	3.9** (2.6; 5.3)	6.5 pp
Combined (2019–2022)	3.8** (2.9; 4.6)	6.2%	0.5 (-0.2; 1.2)	0.7%	3.3** (2.3; 4.2)	5.5 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	0.5 (-0.3; 1.3)	0.8 pp	0.2 (-0.4; 0.7)	0.8 pp	0.4 (-0.6; 1.3)	0.7 pp
2020	1.0 (-0.2; 2.1)	2.0 pp	0 (-0.7; 0.7)	2.0 pp	1.0 (-0.2; 2.1)	2.0 pp
2021	1.3** (0.4; 2.3)	2.3 pp	-0.4 (-0.9; 0.2)	2.3 pp	1.7** (0.5; 2.8)	2.8 pp
2022	1.7** (0.6; 2.8)	2.9 pp	-0.1 (-0.9; 0.7)	2.9 pp	1.8** (0.4; 3.2)	3.0 pp
Combined (2019–2022)	1.0** (0.3; 1.8)	1.9 pp	-0.1 (-0.5; 0.4)	1.9 pp	1.1** (0.2; 2.0)	2.0 pp

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

* $p < 0.10$; ** $p < 0.05$

CI = confidence interval; MDAPM = Maryland All-Payer Model; pp = percentage point.

Exhibit D.9. Impacts of the Maryland Model on quality-of-care outcomes by place

	Difference-in-differences impact estimate for high SVI		Difference-in-differences impact estimate for low SVI		Triple-difference impact estimate for high – low SVI	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
Potentially preventable admissions (number per 1,000 beneficiaries per year)						
Early MDAPM period						
2014	-3.7* (-7.1; -0.3)	-4.5%	0.4 (-0.9; 1.6)	1.0%	-4.0* (-7.6; -0.5)	-5.5 pp
2015	-3.7* (-7.0; -0.4)	-4.7%	0.9 (-0.3; 2.1)	2.2%	-4.6** (-7.9; -1.4)	-6.9 pp
2016	-2.8 (-5.8; 0.2)	-3.7%	0.1 (-1.2; 1.4)	0.3%	-2.9 (-5.9; 0)	-4.0 pp
Later MDAPM period						
2017	-9.5** (-12.9; -6.0)	-11.9%	-1.8* (-3.3; -0.2)	-4.3%	-7.7** (-11.2; -4.1)	-7.6 pp
2018	-12.5** (-16.3; -8.8)	-15.8%	-4.0** (-6.4; -1.7)	-10.0%	-8.5** (-12.2; -4.8)	-5.8 pp
Combined (2017–2018)	-11** (-14.4; -7.6)	-13.9%	-2.9** (-4.7; -1)	-7.1%	-8.1** (-11.5; -4.7)	-6.8 pp
MD TCOC period						
2019	-13.4** (-17.1; -9.7)	-17.1%	-5.8** (-8.0; -3.6)	-14.4%	-7.6** (-11.4; -3.8)	-2.7 pp
2020	-13.4** (-16.8; -10.0)	-21.2%	-4.3** (-6.3; -2.4)	-14.2%	-9.0** (-12.5; -5.6)	-7.0 pp
2021	-12.7** (-16.8; -8.5)	-21.0%	-5.0** (-7.1; -2.8)	-16.3%	-7.7** (-11.5; -3.9)	-4.7 pp
2022	-10.3** (-14.8; -5.9)	-17.3%	-4.9** (-7.2; -2.7)	-15.9%	-5.4** (-9.6; -1.2)	-1.4 pp
Combined (2019–2022)	-12.5** (-16.2; -8.9)	-19.1%	-5.0** (-7.1; -3.0)	-15.2%	-7.5** (-11; -4.0)	-3.9 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	-2.4* (-4.5; -0.3)	-3.2 pp	-2.9** (-3.9; -1.9)	-7.3 pp	0.5 (-1.6; 2.7)	4.1 pp
2020	-2.3 (-5.0; 0.3)	-7.3 pp	-1.4* (-2.7; -0.2)	-7.1 pp	-0.9 (-3.6; 1.8)	-0.2 pp
2021	-1.6 (-4.9; 1.6)	-7.1 pp	-2.1** (-3.3; -0.8)	-9.2 pp	0.4 (-2.8; 3.6)	2.1 pp
2022	0.7 (-2.4; 3.7)	-3.4 pp	-2** (-3.4; -0.7)	-8.8 pp	2.7 (-0.3; 5.7)	5.4 pp
Combined (2019–2022)	-1.5 (-3.8; 0.9)	-5.2 pp	-2.1** (-3.2; -1.1)	-8.1 pp	0.6 (-1.7; 3)	2.9 pp

	Difference-in-differences impact estimate for high SVI		Difference-in-differences impact estimate for low SVI		Triple-difference impact estimate for high – low SVI	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
30-day post-discharge unplanned readmissions (percentage of discharges)						
Early MDAPM period						
2014	-0.8** (-1.3; -0.2)	-3.7%	-0.1 (-0.4; 0.2)	-0.6%	-0.7* (-1.3; 0)	-3.1 pp
2015	-1.4** (-2.0; -0.8)	-6.7%	-0.5** (-0.8; -0.2)	-3.1%	-0.9** (-1.6; -0.3)	-3.6 pp
2016	-1.8** (-2.4; -1.2)	-8.7%	-0.8** (-1.1; -0.6)	-5.0%	-1.0** (-1.6; -0.3)	-3.7 pp
Later MDAPM period						
2017	-2.2** (-2.9; -1.5)	-10.5%	-1.0** (-1.4; -0.7)	-6.2%	-1.1** (-1.9; -0.4)	-4.3 pp
2018	-2.2** (-3.0; -1.4)	-10.6%	-1.0** (-1.3; -0.7)	-6.2%	-1.2** (-2; -0.4)	-4.4 pp
Combined (2017–2018)	-2.2** (-2.9; -1.5)	-10.6%	-1.0** (-1.3; -0.8)	-6.2%	-1.2** (-1.8; -0.5)	-4.4 pp
MD TCOC period						
2019	-3.0** (-3.7; -2.3)	-14.4%	-1.2** (-1.6; -0.9)	-7.4%	-1.8** (-2.5; -1.0)	-7.0 pp
2020	-3.1** (-3.8; -2.4)	-14.8%	-1.4** (-1.8; -1.0)	-8.6%	-1.7** (-2.5; -0.9)	-6.2 pp
2021	-2.1** (-2.8; -1.5)	-10.1%	-0.9** (-1.2; -0.6)	-5.5%	-1.2** (-1.9; -0.6)	-4.6 pp
2022	-2.0** (-2.9; -1.0)	-9.6%	-1.2** (-1.5; -0.8)	-7.4%	-0.8 (-1.8; 0.2)	-2.2 pp
Combined (2019–2022)	-2.6** (-3.2; -1.9)	-12.5%	-1.2** (-1.4; -0.9)	-7.4%	-1.4** (-2.1; -0.8)	-5.1 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	-0.8** (-1.3; -0.3)	-3.8 pp	-0.2 (-0.5; 0.1)	-1.2 pp	-0.6* (-1.2; 0)	-2.6 pp
2020	-0.9** (-1.4; -0.4)	-4.2 pp	-0.4 (-0.8; 0)	-2.4 pp	-0.5 (-1.2; 0.1)	-1.8 pp
2021	0.1 (-0.5; 0.7)	0.5 pp	0.2 (-0.2; 0.5)	0.7 pp	-0.1 (-0.7; 0.6)	-0.2 pp
2022	0.2 (-0.5; 0.9)	1.0 pp	-0.1 (-0.5; 0.2)	-1.2 pp	0.4 (-0.5; 1.2)	2.2 pp
Combined (2019–2022)	-0.4 (-0.8; 0)	-1.9 pp	-0.1 (-0.4; 0.1)	-1.2 pp	-0.3 (-0.7; 0.2)	-0.7 pp
Timely follow-up after acute exacerbation of chronic conditions (percentage of discharges)						
Early MDAPM period						
2014	0 (-1.0; 0.9)	-0.0%	0.1 (-0.7; 0.8)	0.1%	-0.1 (-1.3; 1.1)	-0.1 pp
2015	1.9** (1.1; 2.8)	3.0%	1.0** (0.3; 1.7)	1.4%	0.9 (-0.2; 2.0)	1.6 pp
2016	2.5** (1.5; 3.5)	3.9%	1.0** (0.3; 1.8)	1.4%	1.4** (0.3; 2.5)	2.5 pp
Later MDAPM period						
2017	1.9** (0.8; 2.9)	2.9%	1.2** (0.5; 2.0)	1.6%	0.6 (-0.6; 1.8)	1.3 pp
2018	2.5** (1.3; 3.7)	3.8%	0.6 (-0.2; 1.4)	0.8%	1.9** (0.5; 3.3)	3.0 pp
Combined (2017–2018)	2.2** (1.2; 3.2)	3.4%	0.9** (0.2; 1.7)	1.2%	1.2* (0.1; 2.4)	2.2 pp

	Difference-in-differences impact estimate for high SVI		Difference-in-differences impact estimate for low SVI		Triple-difference impact estimate for high – low SVI	
	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	% Impact[a]	Estimate (90% CI)	Difference in % impacts (in percentage points)
MD TCOC period						
2019	2.4** (1.2; 3.5)	3.7%	1.2** (0.4; 2.1)	1.6%	1.1 (-0.2; 2.5)	2.1 pp
2020	1.9** (0.7; 3.1)	3.1%	1.6** (0.5; 2.6)	2.3%	0.3 (-1.1; 1.7)	0.8 pp
2021	2.8** (1.8; 3.8)	4.4%	0.6 (-0.5; 1.6)	0.8%	2.2** (0.7; 3.7)	3.6 pp
2022	3.4** (2.2; 4.5)	5.3%	1.0* (0; 2.0)	1.4%	2.3** (1.0; 3.7)	3.9 pp
Combined (2019–2022)	2.5** (1.6; 3.5)	3.9%	1.1** (0.3; 2.0)	1.5%	1.4** (0.3; 2.6)	2.4 pp
Difference in estimates during MD TCOC period and later MDAPM period						
2019	0.2 (-0.5; 0.9)	0.3 pp	0.3 (-0.2; 0.9)	0.4 pp	-0.1 (-1.0; 0.8)	-0.1 pp
2020	-0.3 (-1.4; 0.8)	-0.3 pp	0.6 (-0.2; 1.4)	1.1 pp	-0.9 (-2.2; 0.3)	-1.4 pp
2021	0.6 (-0.5; 1.7)	1.0 pp	-0.3 (-1.1; 0.4)	-0.4 pp	1.0 (-0.5; 2.4)	1.4 pp
2022	1.2* (0.1; 2.2)	1.9 pp	0.1 (-0.7; 0.9)	0.2 pp	1.1 (-0.1; 2.3)	1.7 pp
Combined (2019–2022)	0.4 (-0.4; 1.1)	0.5 pp	0.2 (-0.3; 0.7)	0.3 pp	0.2 (-0.7; 1.1)	0.2 pp

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year in Maryland minus the difference-in-differences impact estimate for the year.

* $p < 0.10$; ** $p < 0.05$

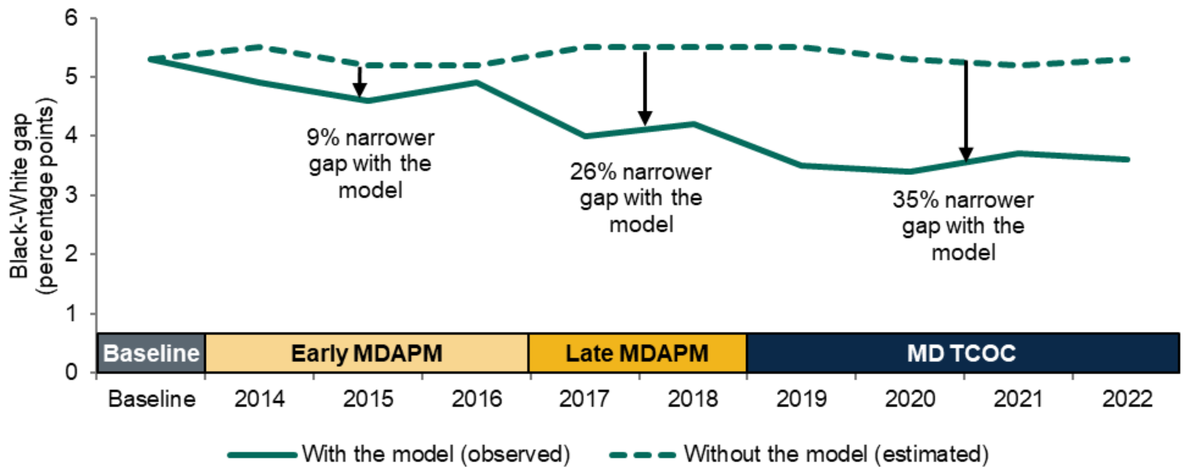
CI = confidence interval; MDAPM = Maryland All-Payer Model; pp= percentage point; SVI = Social Vulnerability Index.

D.4.2. Additional figures showing changes in quality gaps over time

In Chapter 4, we presented figures showing changes in quality gaps over time for potentially preventable admissions. In this section, we present the same figures for the other two quality outcomes—30-day unplanned readmissions and timely follow-up. Using Exhibit D.10 as an example, the following is a description of how readers can interpret these figures:

- At baseline (2011–2013), readmission rates were 5.3 percentage points higher for Black beneficiaries in Maryland relative to White beneficiaries.
- On average, across the early MDAPM period (2014–2016), the observed readmission rates were 4.8 percentage points higher for Black beneficiaries—with yearly differences shown as the black solid line. We calculated what the difference in these readmission rates would have been in the absence of the model (see Section E.3.2 for more details)— as the black dashed line— by adding the estimated impact of the model during the early MDAPM period (the weighted average of the triple-difference impact estimates during the early MDAPM period in Exhibit D.8) to the observed difference in 30-day unplanned readmission rates. The dashboard line averages 5.3 percentage points across the early MDAPM period, showing a change in the Black–White gap of 0.5 percentage points (5.3 – 4.8) or 9% (0.5 / 5.3).
- The dashed line represents the estimated disparity during each period in the absence of the model and can increase or decrease from baseline due to (a) changes in disparities in the comparison group over time or (b) shifting demographic composition in Maryland. However, in Exhibit D.10, the dashed line remains relatively stable over time.
- We used the same calculations for the late MDAPM period and MD TCOC periods.

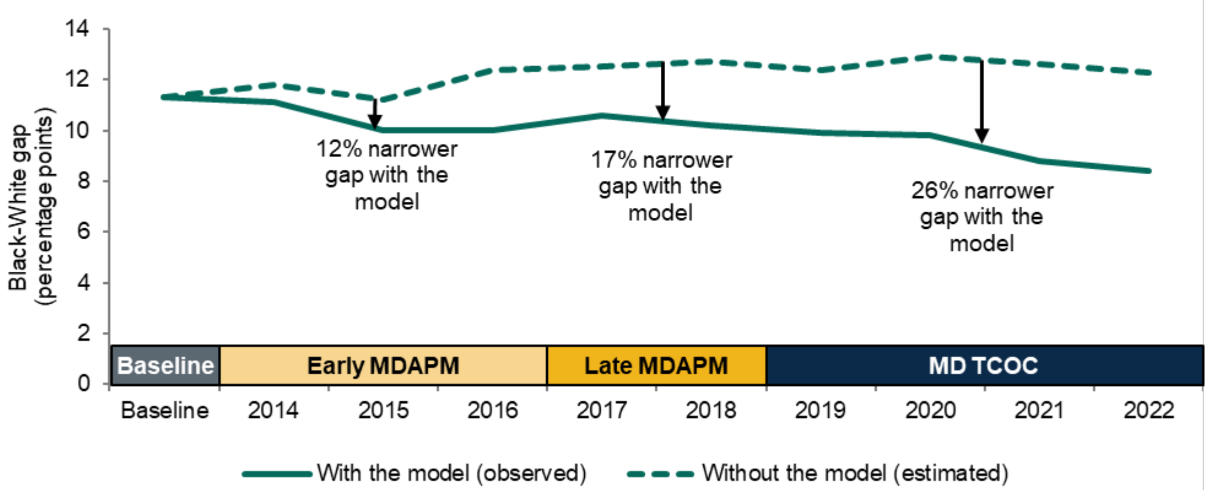
Exhibit D.10. Most of the narrowing of Black–White gaps in 30-day unplanned readmissions observed during the Maryland Model occurred during the MDAPM period



Note: This figure shows the observed (actual, unadjusted) difference, or gap, in rates of 30-day unplanned readmissions for Black and White beneficiaries in Maryland as the black solid line. It also shows what we estimate the Black–White gap would have been in the absence of the Maryland Model (dashed line). To calculate that estimated gap, we first estimated the 30-day unplanned readmission rate for each racial group in each year (equal to the unadjusted rate for that year minus the impact estimate for that year). We then took the difference of those estimated rates for Black and White beneficiaries in each year.

MDAPM = Maryland All-Payer Model.

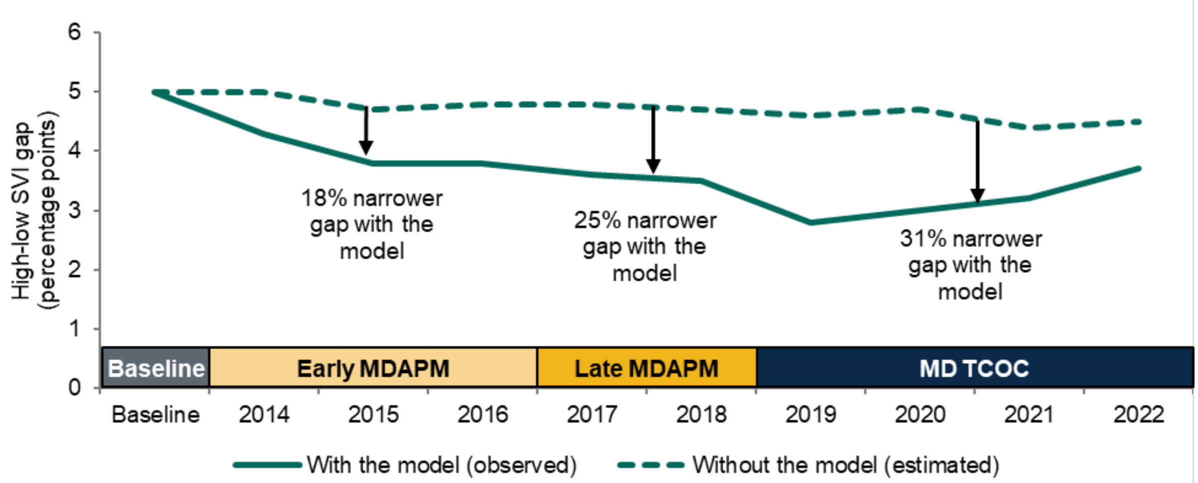
Exhibit D.11. Most of the narrowing of Black–White gaps in timely follow-up observed during the Maryland Model occurred during the MDAPM period



Note: See the notes in Exhibit D.10

MDAPM = Maryland All-Payer Model.

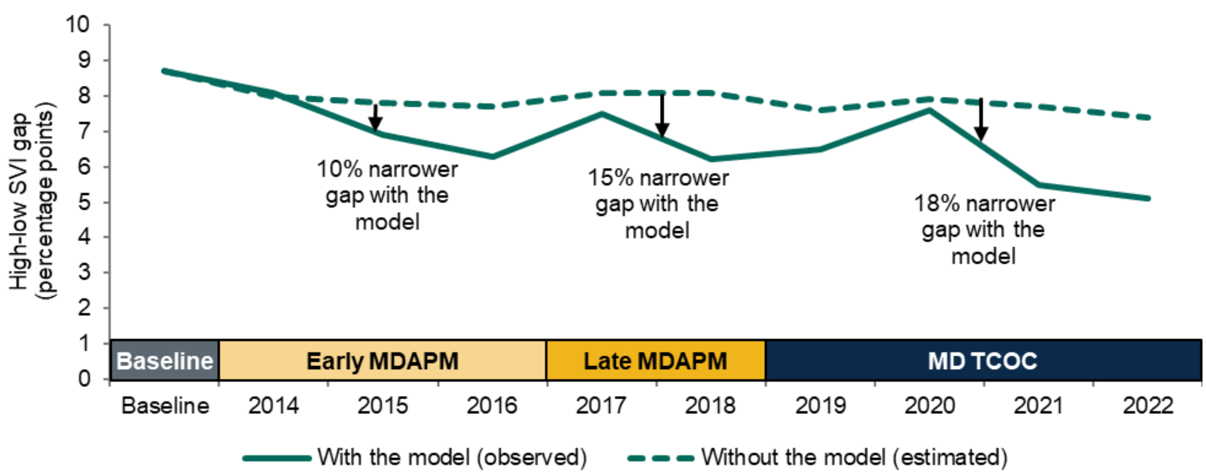
Exhibit D.12. Most of the narrowing of high–low SVI gaps in 30-day unplanned readmissions observed during the Maryland Model occurred during the MDAPM period



Note: See the notes in Exhibit D.10.

MDAPM = Maryland All-Payer Model; SVI = Social Vulnerability Index.

Exhibit D.13. Most of the narrowing of high–low SVI gaps in timely follow-up observed during the Maryland Model occurred during the MDAPM period



Note: See the notes in Exhibit D.10.

MDAPM = Maryland All-Payer Model; SVI = Social Vulnerability Index.

D.4.3. Detailed exploratory analysis of shifts in the location of care by race and place

This section contains the results of an exploratory analysis designed to better understand the impacts of shifts from inpatient admissions to observation stays and community discharges on changes in quality gaps. As described in Chapter 4, these shifts in the site of care could reflect either benefits or harms to beneficiaries. In this exploratory analysis, we first calculated the impacts of the model on inpatient admissions by race and place. We then calculated impacts of the model, by race and place, on emergency department visits that ended in (a) observation stays or (b) community discharges. Finally, we used these estimates to calculate the proportion of all impacts on inpatient admissions attributable to impacts on observation stays or community discharges. The following is a description of how readers can interpret Exhibit D.14:

- For Black beneficiaries, the model led to 89 fewer admissions per 1,000 beneficiaries per year, but we estimate that 26 of those admissions were shifted to observation stays and 9 of those admissions were shifted to community discharges. Thus, we estimate that 30% of the impacts on inpatient admissions for Black beneficiaries (26 / 89) were attributable to shifts towards observation stays, 10% (9 / 89) were attributable to shifts towards community discharges, and therefore 40% ([26 + 9] / 89) were due to the combination of shifts towards observation stays and community discharges.
- Impacts on inpatient admissions were smaller for White beneficiaries (a reduction of 30 admissions per 1,000 beneficiaries per year). The model also had a smaller impact on observation stays (an increase of 12 visits per 1,000 beneficiaries per year) for White compared to Black beneficiaries. However, like Black beneficiaries, about one-third of the reductions in inpatient admissions reflected shifts to observation stays (38%). There were no impacts on community discharges for White beneficiaries.
- The first two rows in the table indicate how much shifts toward observation stays or community discharges explain the impacts on inpatient admissions within subgroups defined by race, but not how shifts in care affect the change in gaps by subgroup— the focus of Chapter 4 which can be found in the "Difference" row. The model reduced the Black–White differences in inpatient admissions by 59 admissions per 1,000 beneficiaries per year[38], but we estimate that 15 of these admissions were the results of shifts to observation stays and 11 to community discharges. Thus, we estimate that 25% of the impacts on Black–White differences in inpatient admissions (15 / 59) were attributable to shifts towards observation stays, 18% (11 / 59) were attributable to shifts towards community discharges, and therefore 43% ([15 + 11] / 59) were due to the combination of shifts towards observation stays and community discharges.
- We calculated the same quantities by place.

These exploratory results should be interpreted with caution, particularly impacts on observation stays. We found diverging trends in observation stays between Maryland and the comparison group at baseline, which violates one of our key assumptions for estimating impacts. Trends suggest that Maryland may have been on a path to moving Black and White beneficiaries to observation before MDAPM.

[38] Based on triple-difference estimates

Exhibit D.14. Estimated proportion of impacts on hospital admissions attributable to impacts on observation stays or community discharges from the ED, 2019–2022

	Impacts per 1,000 beneficiaries per year during MD TCOC (2019–2022)			Proportion of impacts on all admissions attributable to shifts towards			Proportion of differences by race and place in impacts on all admissions attributable to shifts towards		
	Impacts on hospital admissions	Impacts on observation stays from the ED[a]	Impacts on community discharges from the ED[a]	Observation stays	Community discharges	Observation stays or community discharges	Observation stays	Community discharges	Observation stays or community discharges
Black	-89**	26**	9**	30%	10%	40%	n.a.	n.a.	n.a.
White	-30**	12**	-1	38%	-5%	34%	n.a.	n.a.	n.a.
Difference	-59**	15**	11**	n.a.	n.a.	n.a.	25%	18%	43%
High SVI	-73**	26**	3	36%	4%	40%	n.a.	n.a.	n.a.
Low SVI	-35**	11**	-1	31%	-3%	29%	n.a.	n.a.	n.a.
Difference	-38**	15**	4	n.a.	n.a.	n.a.	40%	11%	51%

Note: An ED visit can either end in a hospital admission, observation stay, or community discharge. Thus, some of the reductions in hospital admissions could be attributed to increases in observation stays or discharges home. For example, we estimate that of the 89 hospital admissions per 1,000 Black beneficiaries per year prevented by the model, 35 (40%) were either additional observation stays (26) or community discharges (9). Similarly, we estimate that of the 59 additional hospital admissions reduced for Black beneficiaries compared to White beneficiaries, 26 (43%) of those admissions were either additional observation stays (15) or community discharges (11).

[a] Modeled at the ED visit level and converted to the scale of per 1,000 beneficiaries per year.

* $p < 0.10$; ** $p < 0.05$

ED = emergency department; n.a. = not applicable; SVI = Social Vulnerability Index.

Appendix E. Methods and Supplemental Results for Estimating the Added Effect of the Maryland Primary Care Program for Medicare Fee-For-Service Beneficiaries

E.1. Design and methods for estimating impacts

E.1.1. Analysis overview

The main impact evaluation for the Maryland Model uses an out-of-state comparison group to estimate the effects of all Maryland Model components combined for Maryland Medicare fee-for-service (FFS) beneficiaries throughout the state. In 2019, CMS and the Maryland Department of Health introduced the Maryland Primary Care Program (MDPCP). This program represents a significant investment of CMS and state resources into primary care transformation in Maryland. To evaluate the impact of these investments, CMS is interested in understanding what the added effect of MDPCP is on key outcomes in the state beyond the impacts of other Maryland Model components (that is, other incentives and supports like hospital global budgets).

We use beneficiary- and discharge-level difference-in-differences models with matched comparison practices drawn from *within Maryland* and practice-level fixed effects to estimate the added effects of MDPCP. Specifically, the impact estimate will be (1) the regression-adjusted change in outcomes for Medicare FFS beneficiaries assigned to MDPCP practices (2019 starters only) from the two-year baseline period (2017–2018) to a four-year intervention period (2019–2022) minus (2) the adjusted, contemporaneous changes for beneficiaries attributed to similar, non-MDPCP practices located in Maryland and matched to look similar to the MDPCP practices. To make the analysis feasible within budget and to maximize the ability to detect any effects that accrue over time, we limited the intervention group to beneficiaries attributed to practices that joined in 2019 and excluded practices that joined MDPCP later (2020–2021) from both the intervention and potential comparison group; however, we did include these practices in attribution to allow them to compete for beneficiaries. The main advantage of this design is that the comparison group of non-MDPCP practices in Maryland would be subject to all MD TCOC incentives and supports (except for MDPCP incentives and supports), and, therefore, effects from other aspects of the model (like global budgets) would be netted out in the difference-in-differences analysis.

E.1.2. Data sources

E.1.2a MDPCP participation data

We obtained MDPCP participation data from The Lewin Group, the contractor that helps implement MDPCP. The data included a roster of the practices participating in MDPCP and the years that they began and withdrew from the program, if applicable. The data also included the practice’s Tax Identification Number (TIN), a list of practitioners working at each practice, a list of Medicare FFS beneficiaries attributed to each practice, and a list of Care Transformation Organizations (CTOs) that supported MDPCP practices.

E.1.2b OneKey™ by IQVIA

We purchased yearly rosters from 2019 to 2022 from IQVIA, a commercial health care data vendor that maintains and verifies lists of practitioners who work in practices throughout the United States. The OneKey™ data contains information about practices, including the practices' location and specialty, the providers affiliated with the practice, and the corporate parents of the practices. It also includes the providers affiliated with the practice (such as name, specialty, and National Provider Identifier [NPI]) and corporate parents of the practices (including ownership type and name).

We opted to use practice and practitioner information from OneKey™ for each year to construct the matching and analytic files and identified MDPCP practices in OneKey™ using the Lewin Group data. This was because, as part of the evaluation research design, we constructed a matched comparison group of practices not participating in MDPCP. For practices not participating in MDPCP, we had to rely on OneKey™ data for all practice and practitioner information. By using OneKey™ data for all practices, we removed the possibility of bias from using different data sources for MDPCP versus non-MDPCP practices.

To link the MDPCP participant data with the OneKey™ data, we merged the NPIs of the affiliated providers at the practices and retained all practices that were found in both data sets. We compared practice zip code, address, and name (using Levenshtein distance, which is the smallest number of edits necessary to change one string into the other) among these overlapping practices. Matches were confirmed if all three variables matched across the two data sets; otherwise, the matches were adjudicated by two independent reviewers. For MDPCP practices not found in the OneKey™ data, we appended practice and practitioner information from those practices' directly from the Lewin Group.

E.1.2c Other secondary data sources for attribution

We used National Plan and Provider Enumeration System (NPPES), a national data set with one record for each health care provider with an NPI and information on the NPI's specialty taxonomy, to identify which NPIs in Maryland were primary care providers (PCPs). We used Medicare Data on Provider and Specialty (MD-PPAS) to assign a TIN to MDPCP and comparison practices. MD-PPAS is a provider-level file that contains information on all NPIs with Medicare Part B claims and provides information on the NPI's affiliated TINs.

E.1.2d Medicare FFS claims data

To identify attributed beneficiaries for MDPCP and comparison practices and to construct the beneficiary- and episode-level analytic files, we used Medicare FFS claims data. Beneficiaries had to meet the eligibility requirements for the main impacts analysis to be included in the MDPCP analysis—that is, be alive and enrolled in Medicare FFS Parts A and B for at least one month of the year, with Medicare as the primary payer.

E.1.3. Analytic file

E.1.3a Intervention group

The intervention group consists of Medicare FFS beneficiaries assigned to 361 MDPCP practices that started in 2019, which is about 95% of the 380 MDPCP practices that started in 2019. We started with the full set of 380 practices that joined MDPCP in 2019 and then applied several restrictions, such as removing practices with fewer than 50 assigned beneficiaries in the baseline period (2017–2018) to both the intervention and comparison groups. For intervention group and comparison group sample sizes, see Exhibit E.14.

E1.3b Comparison group

The comparison group consists of Medicare FFS beneficiaries assigned to 450 group practices that provide primary care in Maryland and never participated in MDPCP from 2019 to 2022 (effective sample size with weighting is 261[39]). To identify the comparison group practices, we applied restrictions to the pool of OneKey™ practices in 2019 and generated matching weights for these practices (see Section E.1.5). The attribution and assignment steps are detailed below.

Exhibit E.1 compares the beneficiary characteristics in 2022 among those attributed to 2019 MDPCP starters, 2020–2021 MDPCP starters, and non-MDPPC beneficiaries with and without a PCP. In this table, characteristics for MDPCP beneficiaries are based on the payment attribution, while the non-MDPCP beneficiary attribution is based on the evaluation attribution. We find that 2019 MDPCP starters are older and less likely to be non-Hispanic Black or dually eligible for Medicaid than 2020–2021 MDPCP starters. Maryland beneficiaries who are not attributed to MDPCP with a PCP have a higher Hierarchical Condition Category (HCC) score than MDPCP beneficiaries, as well as higher utilization and spending. Maryland beneficiaries who are not attributed to MDPCP and are not attributed to a PCP have a lower HCC score, utilization, and spending than those attributed to a PCP, including both MDPCP and non-MDPCP beneficiaries.

[39] Effective sample size is the number of practices that, if weighted equally, would have the same amount of precision as the weighted sample size.

Exhibit E.1. Comparison of beneficiary characteristics in 2022 across 2019 MDPCP practices, 2020–2022 MDPCP practices, non-MDPCP practices, and unattributed beneficiaries

Variable description	Attributed to MDPCP 2019 starters	Attributed to MDPCP 2020–2021 starters	Non-MDPCP Maryland FFS beneficiaries in 2022 with a PCP	Non-MDPCP Maryland FFS beneficiaries in 2022 without a PCP
Number of FFS Medicare beneficiaries	304,755	96,063	247,722	130,477
Average HCC score	1.14	1.11	1.34	0.77
Percent living in rural areas	17.7%	16.0%	17.5%	14.1%
Average Social Vulnerability Index	0.35	0.37	0.37	0.38
Average age	73.2	71.8	72.5	67.8
Percent female	59.4%	59.4%	58.2%	48.0%
Percent non-Hispanic Black	18.8%	27.3%	24.1%	30.3%
Percent dually eligible for Medicare	12.8%	19.7%	21.2%	22.3%
Percent with an original reason for Medicare entitlement other than aged (e.g., disability or ESRD)	15.4%	19.9%	21.2%	19.6%
Mean all-cause admissions per 1,000 beneficiaries	263	270	373	169
Mean all-cause ED and observation visits per 1,000 visits	391	422	487	258
Mean total Medicare FFS spending per beneficiary per year	16,319	16,423	22,486	10,459

ED = emergency department; ESRD = end-stage renal disease; FFS= fee for service; HCC = Hierarchical Condition Category; MDPCP = Maryland Primary Care Program; PCP = primary care provider.

E.1.4. Attribution and assignment

In this section, we describe the process for beneficiary attribution and assignment for the evaluation of the MDPCP. This process consisted of five steps:

- First, we identified the set of primary care practices in Maryland that could compete for beneficiaries in the attribution and assignment process.
- Second, we assigned TINs to each primary care practice in our data (assigning TINs at the practice-level was necessary to attribute beneficiaries to practices).
- Third, we identified eligible Medicare claims data needed for attribution.
- Fourth, we attributed eligible Medicare beneficiaries to a single primary care practice in each year based either on their most recent visits for certain services, such as Medicare Annual Wellness Visit, or the plurality of their primary care services in the previous two years.

- Fifth, we assigned beneficiaries to practices during the baseline and performance years to the first practice to which they were attributed in each period.

We describe each of the above steps in more detail below. Our attribution approach is similar, but not identical to, the process where beneficiaries are attributed to MDPCP practices for the purpose of calculating model payments. Therefore, we conclude by comparing how our evaluation attribution process differs from the process of attributing beneficiaries for payment.

E.1.4a. Attribution and assignment steps

Step 1: Identify primary care practices. We define each practice in Maryland for attribution as comprising a unique group of practitioners who work at the address at a given point. We used the yearly rosters from 2019 to 2022 from IQVIA. The IQVIA OneKey™ database starts in 2019, but components of our attribution approach and evaluation design (namely incorporating a two-year baseline period) required we had practice composition data for 2015, 2016, 2017, and 2018 (because there is a 24-month lookback period for attribution). Practice–provider data like OneKey™ produced by a different vendor exist for years prior to 2019, but an examination of these data in 2018 revealed discrepancies in practice compositions to the OneKey™ data in 2019 that were greater than expected based on a difference of a single year. Rather than combine two meaningfully different sets of data, we decided that 2019 OneKey™ data would reflect practice composition for all years prior to 2019. This decision effectively assumes stable practice compositions for years prior to 2019. We also conducted an analysis of OneKey™ data in 2019 and later years to examine differences in practice compositions year to year. We found that practice compositions were relatively stable; for example, less than 10% of all providers in a practice in 2019 were in a different practice in 2020. However, we do see an increase in the number of comparison practices in 2021 and 2022 relative to 2020 (see Exhibit E.4), which may either reflect increased data completeness over time or newly established practices in Maryland.

Step 2. Assign yearly TINs to each primary care practice. OneKey™ data do not include TINs, which are necessary for beneficiary attribution to practices in later steps. Therefore, we used MD-PPAS to first assign TINs to each *provider* in the practice with an eligible primary specialty taxonomy code in NPPES (Exhibit E.2) and then assigned one TIN per year to the entire practice based on the plurality of Medicare line items billed to a TIN by eligible PCPs within each practice. If there was a tie among the most common number of line items billed to a TIN, we assigned the practice TIN based on the TIN that corresponded to the plurality of allowed charges billed at the practice among eligible providers.

Exhibit E.2. Eligible primary care specialty codes for TIN assignment

Specialty description	Taxonomy code
Family Medicine	207Q00000X
Adolescent Medicine	207QA0000X, 207R00000X
Addiction Medicine	207QA0401X, 207RA0401X
Adult Medicine	207QA0505X
Geriatric Medicine	207QG0300X, 207RG0300X
Hospice and Palliative Medicine	207QH0002X, 207RH0002X, 207VH0002X
Internal Medicine	207R00000X
Obstetrics & Gynecology	207V00000X, 363LX0001X
Gynecology	207VG0400X
Maternal & Fetal Medicine	207VM0101X
Obstetrics	207VX0000X
Pediatrics	208000000X, 363LP0200X, 364SP0200X
Psychiatry & Neurology: Psychiatry	2084P0800X
General Practice	208D00000X
Physician Assistant	363A00000X
Medical	363AM0700X
Nurse Practitioner	363L00000X
Acute Care	363LA2100X, 364SA2100X
Adult Health	363LA2200X, 364SA2200X
Community Health	363LC1500X
Family	363LF0000X
Gerontology	363LG0600X, 364SG0600X
Psych/Mental Health	363LP0808X, 364SP0808X
Perinatal	363LP1700X, 364SP1700X
Primary Care	363LP2300X
School	363LS0200X
Women's Health	363LW0102X, 364SW0102X
Clinical Nurse Specialist	364S00000X
Community Health/Public Health	364SC1501X
Chronic Care	364SC2300X
Family Health	364SF0001X
Home Health	364SH0200X
Medical-Surgical	364SM0705X
Psych/Mental Health, Child & Adolescent [a]	364SP0807X
Psych/Mental Health, Adult [a]	364SP0809X
Psych/Mental Health, Child & Family [a]	364SP0810X
Psych/Mental Health, Chronically Ill [a]	364SP0811X
Psych/Mental Health, Community [a]	364SP0812X
Psych/Mental Health, Geropsychiatric [a]	364SP0813X
Pediatrics: Adolescent Medicine	2080A0000X
Pediatrics: Hospice and Palliative Medicine	2080H0002X

Specialty description	Taxonomy code
Physician Assistants & Advanced Practice Nursing Providers Advanced Practice Midwife	367A00000X
Ambulatory Health Care Facilities Clinic: Center Primary Care	261QP2300X
Preventive Health: Public Health & General Preventive Medicine	2083P0901X

Notes: [a] Following MDPCP payment attribution criteria, we required providers with these primary specialties be co-located within a practice that also had a practitioner with another eligible specialty (not psych/mental health related) to be considered eligible.

If no primary specialty for the provider was listed in NPPES, then we required all listed *secondary* specialties be eligible for the provider to be eligible (following MDPCP payment attribution criteria).

MDPCP = Maryland Primary Care Program; NPPES = National Plan and Provider Enumeration System; TIN = Tax Identification Number.

We assign a single TIN to a practice during a year, but it was possible that a practice's TIN could change within a year, which might show up as providers billing to two or more TINs within the same year. To account for this change, we implemented a two-step process to our practice-provider data. First, we identified all practices that had an assigned TIN switch over two years. Second, we created duplicate records for all providers in practices that switched TINs for the year prior to the TIN change.

Prior to running beneficiary attribution, we also required that a given NPI-TIN combination be associated with a single practice per year. For cases where a given NPI-TIN combination was associated with more than one practice in the same year, we implemented the following decision rules:

1. If a given NPI-TIN combination appeared in both a MDPCP 2019 starter and a potential comparison practice, we kept only the NPI-TIN combination in the 2019 MDPCP practice.
2. If a given NPI-TIN combination appeared in both a MDPCP practice that started *after* 2019 and a potential comparison practice, we kept only the NPI-TIN combination in the MDPCP practice that started after 2019.[40]
3. If a given NPI-TIN combination appeared in both a MDPCP practice that started in 2019 and a MDPCP practice that started after 2019, we kept only the NPI-TIN combination in the MDPCP practice that started in 2019.
4. If a given NPI-TIN combination appeared in *multiple* MDPCP practices that started in 2019 (or, separately, multiple potential comparison practices), we randomly chose which practice to keep the NPI-TIN combination retain the randomly chosen NPI-TIN combination with the same practice in all future years (starting in 2015) where we also have cases where that same NPI-TIN appears associated with two practices.

Overall, **over 97%** of TINs we assigned to MDPCP practices matched practice TINs reported in data from the Lewin Group. This is evidence that our process was able to accurately assign the correct TIN to the practices in our sample.

[40] MDPCP practices that started after 2019 are excluded from the evaluation; however, we allow beneficiaries to be attributed to these practices.

Step 3. Identify Medicare claims lines for attribution. We required a Medicare claims line meet four criteria to be included in the attribution process for a given year: claim type, claim date, service type, and specialty of the practitioner who provided the service.

1. **Claim type.** We used national Medicare FFS physician (Part B carrier) and outpatient claims for years 2015 through 2021. We limited claims data to beneficiaries who were enrolled in Medicare Parts A and B at the start of an attribution year and whose primary residence is within a Maryland zip code or a Maryland hospital primary service area zip code.
2. **Claim date.** For each attribution year, our lookback period for claims data was the 24-month period that ended the day before the start of the year. For example, for 2017, the lookback period for claims runs from January 1, 2015, through December 31, 2016.
3. **Service type.** We limit claims data to eligible primary care services using the Healthcare Common Procedure Coding System (HCPCS) codes reported on each claim line. Exhibit E.3 lists the HCPCS codes of services we consider to be related to primary care, which follows the list used for MDPCP payment attribution.[41]

Exhibit E.3. Eligibility for primary care service claims was based on HCPCS codes

Service	HCPCS codes
Office/outpatient visit evaluation and management	99201 -99205 99211-99215
Home Care	99324-99328 99334-99337 99339-99345 99347-99350
Welcome to Medicare	G0402
Annual Wellness Visit	G0438, G0439
Advance care planning	99497
Collaborative care model	G0502-G0504
Cognition and functional assessment for patient with cognitive impairment	G0505
Transitional care management services	99495-99496
Chronic care management services	99490, 99491
Complex chronic care management services	99487, 99489
Assessment/care planning for patients requiring chronic care management services	G0506
Non-complex chronic care management clinical staff time	G2058, 99439
Chronic care management services for a single high-risk disease	G2064, G2065
Care management services for behavioral health conditions	G0507
Cervical/vaginal cancer screening; pelvic and clinical breast exam	G0101
Administration of influenza virus vaccine	G0008
Influenza virus vaccine, split virus, preservative free, enhanced	90662
Influenza virus vaccine, quadrivalent, split virus, preservative free	90686
Influenza virus vaccine, quadrivalent, split virus, when administered	90688

[41] We counted the number of unique (eligible) line items as separate primary care visits—but only allowed multiple line items for the same TIN-NPI combination and service date to count as *one* visit.

Service	HCPCS codes
Influenza virus vaccine, trivalent, split virus, preservative free, when	90656
Influenza assay w/optic	87804
Influenza immunization was ordered or administered	G8482
Influenza virus vaccine, derived from cell cultures, subunit, and preservative	90661
AG0immunofluor; influenza B	87275
Influenza AG, DFA	87276
Influenza virus vaccine, trivalent, split virus, when administered to	90658
Administration of pneumococcal vaccine	G0009
Pneumococcal conjugate vaccine, 13 valent, for intramuscular use	90670
Pneumococcal vaccine, polyvalent	90732
Established patient periodic preventive medicine examination, age 65 years	99397
Preventive medicine services	99385-99387, 99401-99402, 99420
Established patient periodic preventive medicine examination age 18-39 years	99395
Established patient periodic preventive medicine examination age 40-64 years	99396
Adult preventive medicine	90750
Hemoglobin; glycosylated	83036
Hemoglobin; glycosylated (A1C) by device cleared by FDA for home use	83037
Glucose, blood, by glucose monitoring device(s)	82962
Prolonged non-face-to-face evaluation and management services	99358-99359
Prolonged Service with Direct Patient Contact	99354-99355
Federally qualified health center visit, new patient	G0466
Federally qualified health center visit, established patient	G0467
Federally qualified health center visit, initial preventive physical examination or annual wellness visit	G0468
Distant site telehealth services Rural Health Clinics or Federally Qualified Health Centers	G2025
Telephone Evaluation and Management Service Provided by a Physician	99441-99443

FDA = Food and Drug Administration; HCPCS = Healthcare Common Procedure Coding System.

4. **Specialty of practitioner who provider service.** We limited claims data to those from services rendered by a practitioner that had a primary specialty listed in Exhibit E.7.

Step 4. Running the attribution algorithm. The attribution algorithm first linked beneficiary claim-lines with the practice–provider data by TIN and rendering NPI and year. This effectively "stamped" a claim-line with a practice identifier that the NPI-TIN combination was associated with in a given year. We also allowed Federally Qualified Health Centers (FQHCs) to compete in attribution (even though we did not have comprehensive provider information for these sites of care in the OneKey™ data) by linking beneficiaries to the CMS Certification Number that appeared on outpatient claim-lines. We then attributed beneficiaries to the practice that submitted the most recent claim-line for the following types of services: chronic care management, Welcome to Medicare, Initial Preventive Physical Examination, and Annual Wellness Visits. All remaining beneficiaries were attributed to the practice that provided the plurality of their primary care services during the two-year lookback period. Any ties, where two or more practices rendered the same number of primary care services for a beneficiary, were broken by examining the most recent claim-line (that is, the beneficiary was attributed to the practice that rendered the most recent claim-line).

Step 5. Assigning beneficiaries to practices. We assigned beneficiaries to practices during the baseline years (that is, before MDPCP began) and, separately, during the performance years, to the first practice to which they were attributed during each period (baseline and performance), allowing beneficiaries to change practice assignment between the baseline and intervention periods. Following an intent-to-treat approach, a beneficiary continued to be assigned to the same practice for the entire period (either baseline or performance), regardless of whether the beneficiary continued to receive care at that practice, as long as they were eligible for attribution for those subsequent years. For example, if a beneficiary switched from receiving care at a MDPCP practice to receiving care at a comparison practice, we continue to assign the beneficiary among the group that might have benefitted from the intervention.

Assignment helps limit bias in our impact estimates due to changing population composition between the MDPCP and comparison groups. For example, the MDPCP incentive structure could have led practices to increase or reduce the need for in-person primary care services rendered to certain beneficiaries, thus differentially changing the *attribution* samples between the intervention and comparison group over time. Exhibit E.4 shows yearly results from the attribution and assignment process among practices (MDPCP starters in 2019 and potential comparisons) that had at least one assigned beneficiary. The total number of attributed beneficiaries each year was substantially higher in the MDPCP group compared to the potential comparison group. MDPCP practices also had, on average, substantially more assigned beneficiaries in each year compared to potential comparison practices.

Exhibit E.4. MDPCP practices starting in 2019 on average had much higher attributed and assigned beneficiaries compared to potential comparison practices in Maryland

Year	Practice category	Number of practices	Average assigned beneficiaries per practice	Total attributed beneficiaries
2017	MDPCP	367	627	230,095
	Potential comparison	820	168	137,787
2018	MDPCP	373	693	250,529
	Potential comparison	871	174	137,958
2019	MDPCP	374	732	273,779
	Potential comparison	816	173	141,218
2020	MDPCP	376	814	298,770
	Potential comparison	852	182	140,888
2021	MDPCP	380	811	281,978
	Potential comparison	982[42]	161	131,202
2022	MDPCP	380	814	275,758
	Potential comparison	1089	147	125,532

Notes: The above table shows yearly results from the attribution and assignment process for 2017–2022 after limiting to practices with at least one assigned beneficiary.

MDPCP = Maryland Primary Care Program.

E.1.4b. Comparing attribution process for evaluation and model payments

Our attribution process for the MDPCP evaluation is similar to the claims-based attribution method used to attribute beneficiaries to practices for MDPCP payments. There were, however, notable differences between the two approaches that we highlight in Exhibit E.5. These differences largely stem from differences in data sources used across both methods and, for the evaluation, making sure that the attribution process for the evaluation does not favor one practice type over another (for example, allowing more attributed beneficiaries to MDPCP practices over non-MDPCP practices).

[42] The increase in number of practices in 2021 and 2022, which are based on attribution data from 2019-2020 and 2020-2021 respectively, likely reflects increased completeness in the OneKey provider data over time. The 2019 OneKey provider data was used for attribution in all years prior to 2021 (see E.1.4a., Step 1 for rationale).

Exhibit E.5. Key differences between beneficiary attribution methods for payment and evaluation

	Payment attribution	Evaluation attribution
Beneficiary eligibility criteria for attribution	Requires the two evaluation attribution eligibility criteria and these additional restrictions: **1.** Not incarcerated **2.** Not admitted to hospice **3.** No ESRD entitlement **4.** No long-term institutional status **5.** Not enrolled in a Medicaid Chronic Health Home **6.** Not enrolled in the Medicare Shared Savings program through a practice other than the MDPCP practice to which the beneficiary is attributed **7.** Not enrolled in any of the following programs: Independence At Home, Comprehensive Primary Care Plus, Comprehensive ESRD Care Model, Next Generation ACO, Direct Contracting (all options), Primary Care First, Kidney Care First and Comprehensive Kidney Care Contracting Graduated, Professional, and Global Models, or any other Innovation Center models with a no-overlaps policy with the MDPCP	**1.** Alive and enrolled in Medicare FFS Parts A and B at the start of the attribution year with Medicare designated at the primary payer **2.** Address of beneficiary primary residence is within a Maryland zip code or a Maryland hospital primary service area zip code.
Frequency of attribution	Quarterly	Yearly
Source for practice and practitioner rosters	MDPCP participation rosters	OneKey™ data for years 2019–2022 (with the assumption that practice compositions in 2015–2018 were the same as 2019)
Criteria used to identify eligible practitioners for attribution	Practitioners in MDPCP rosters and those with NPPES primary or secondary not in rosters (Exhibit E.2.1)	Practitioners affiliated with OneKey practices, as well as those only in claims data, all restricted to those with NPPES primary or secondary specialty of primary care (Exhibit E.2.1)
Source for TINs	MDPCP participation rosters	Assigned TIN based on MD-PPAS and NPPES data linked to practitioners affiliated with practices in OneKey™
Lookback period for claims	Two years ending four months before the start of the quarter	Two years ending the day before the start of the year
Practices and practitioners with whom MDPCP practices compete for beneficiaries	NPI-TIN combinations grouped within MDPCP practices in model rosters; NPI-TIN combinations not in MDPCP rosters but observed in claims	NPI-TIN combinations grouped according to OneKey™ with an assigned TIN in practices; NPI-TIN combinations not in OneKey™ but observed in claims

	Payment attribution	Evaluation attribution
Attribution algorithm	Attributed based on the following hierarchy: **1.** Looking across the eligible claims for each beneficiary to identify (a) the TIN that appears the most often and (b) the TIN and NPI that appears on the most recent claim. If these match to the TIN-NPI combination from a MDPCP Practice, the beneficiary is attributed to that practice. **2.** All remaining unattributed beneficiaries with claims for CCM, WTM, IPPE, or AWV are attributed to the practice that submitted the most recent claim for CCM, WTM, IPPE, or AWV, regardless of whether the practice is a MDPCP practice. **3.** All remaining unattributed beneficiaries are attributed to the practice that provided the plurality of their primary care services during the lookback period. Any ties, where two or more practices have rendered the same number of primary care services, are broken by examining the most recent claim.	Attribution follows steps 2 and 3 of the payment attribution.

Notes: The above table describes key differences between the MDPCP evaluation attribution process and the attribution process for MDPCP model payments.

ACO = Accountable Care Organization; AWV = Annual Wellness Visit; CCM = chronic care management; ESRD = end-stage renal disease; FFS = fee for service; HCPCS = Healthcare Common Procedure Coding System; IPPE = Initial Preventative Physical Examination; MDPCP = Maryland Primary Care Program; MD-PPAS = Medicare Data on Provider Practice and Specialty; NPI = National Provider Identifier; NPPES = National Plan and Provider Enumeration System; TIN = Taxpayer Identifier Number; WTM = Welcome to Medicare.

E.1.4c. Comparison between attribution sample used for payment and attribution sample used for the evaluation

The analytic sample used for evaluation is not the same as the attributed sample used for the purpose of identifying payments to practices for a variety of reasons. One of the biggest reasons is that attribution for the purpose of payment was only done for MDPCP practices. To limit bias in our impact estimates, we apply the same attribution rules to both MDPCP and potential comparison practices, which means we need to use our own attribution but mirror payment attribution as closely as possible. Exhibit E.6 shows the overlap in number of beneficiaries attributed to 2019 MDPCP practices using payment attribution (based on Lewin Group data) versus evaluation attribution. In general, there was more than 75% overlap in 2020 to 2022 with less overlap (62%) in 2019. The payment attribution methodology was updated in 2020 to attribute more beneficiaries. For consistency, our evaluation approach chose to apply the 2020+ payment methodology to all years in the analysis (2017–2022) and to use the 2019 practitioners in the OneKey data for attribution in 2017–2018 (detailed above), both of which may have led to lower rates of overlap in 2019. There were approximately 10 to 15% of beneficiaries who were in the payment attribution but not the evaluation attribution. This may represent instances where either the NPI lists may differ between OneKey™ data and MDPCP program data. There were also instances where beneficiaries were attributed to both 2019 MDPCP practices and 2020–

2021 MDPCP practices in a given year, because MDPCP attribution is completed on a quarterly basis. Therefore, some of the beneficiaries in only the payment attribution may represent beneficiaries who were attributed to both 2019 and 2020–2021 MDPCP practices for payment but were only attributed to 2020–2021 MDPCP practices in the evaluation attribution. Exhibit E.5 provides the full details on differences between the two attribution approaches.

Exhibit E.6. Beneficiary-level overlap between payment attribution and evaluation attribution for MDPCP beneficiaries in 2019 practices

Attribution year	In payment and evaluation attribution N (%)	Only in evaluation attribution N (%)	Only in payment attribution N (%)
2019	187,006 (62%)	86,773 (29%)	28,397 (9%)
2020	279,499 (81%)	19,271 (6%)	46,872 (14%)
2021	257,803 (77%)	24,175 (7%)	50,731 (15%)
2022	248,484 (76%)	27,274 (8%)	53,318 (16%)

MDPCP = Maryland Primary Care Program.

E.1.5. Developing the matched comparison group

We developed the matched comparison group in four steps:

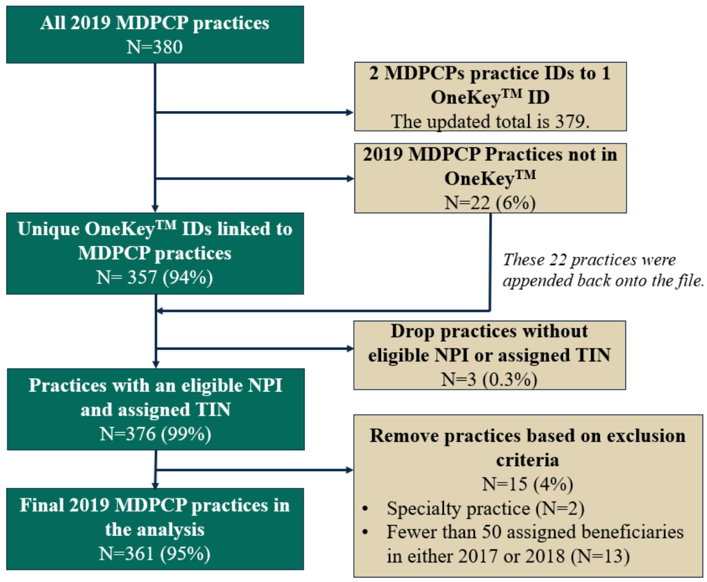
1. Implemented preliminary practice exclusions
2. Identified variables to match on and set criteria for what counts as sufficient balance
3. Used a reweighting method to create the matched comparison group
4. Assessed the quality of the matched comparison group in terms of size, geographic spread, health system composition, balance, and statistical power

In the following sections, we describe each of these four steps. When we developed the comparison group, we explored many alternatives reflecting tradeoffs in different dimensions of quality for the comparison group. We discussed these alternatives with CMS and decided on a final comparison group that we agreed achieved the best balance on the various dimensions. In this section, we report only the results for the final selected comparison group.

E.1.5a. Preliminary practice exclusions

Matching was done at the practice level. We included all non-MDPCP primary care practices in Maryland that met basic criteria in the comparison group. The exclusion criteria, designed to ensure the practices were reasonable matches and we could identify their Medicare patients in claims, included the following: (1) practices with no affiliated NPI that was an MDPCP-eligible primary care provider; (2) practices with no assigned TIN; (3) practices that have fewer than 50 assigned Medicare FFS beneficiaries in 2017 or 2018[43]; (4) practices that are health departments, Veterans Health Administration, or specialty practices[44]; (5) practices that are extreme outliers on key matching variables; and (6) practices with missing information on key outcomes. We also applied these restrictions to the intervention practices to remain consistent across the treatment and comparison groups and not introduce bias. There were 380 MDPCP practices and 1,443 comparison practices in the starting pool, after removing the practices starting MDPCP in 2020–2022 from the OneKey practice pool. After restrictions, there were 361 MDPCP practices and 450 comparison practices remaining (see Exhibits E.7 and E.8 for details on restrictions).

Exhibit E.7. Flowchart of restrictions to MDPCP practices



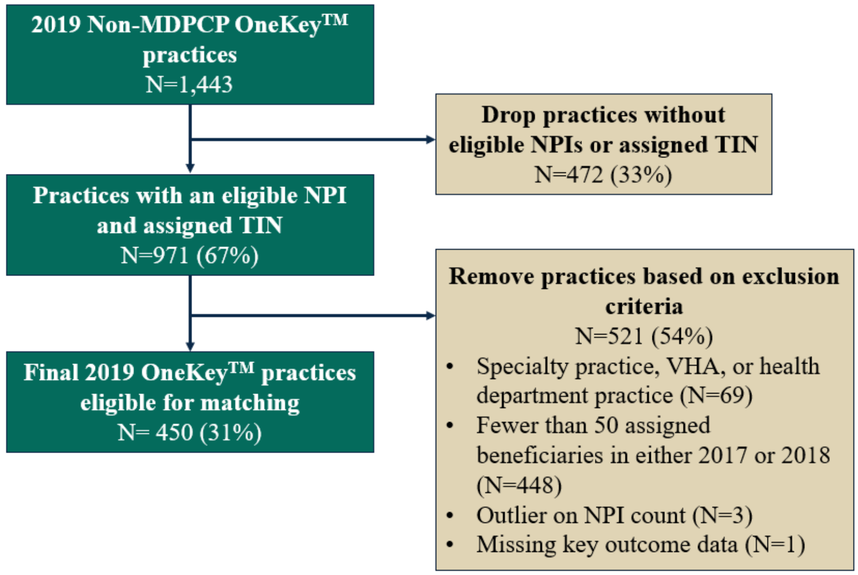
Note: The percentages in the left column (green) refer to the percentage remaining of the total initial practice count, and the percentages in the right column (cream) refer to the percentage removed from the prior step because of the restriction.

MDPCP = Maryland Primary Care Program; NPI = National Provider Identifier; TIN = Tax Identification Number.

[43] We relaxed the MDPCP eligibility requirement from at least 125 beneficiaries to at least 50 beneficiaries to increase the number of practices in the potential comparison pool and to maximize the number of MDPCP practices retained in the analysis. Further, several of the MDPCP practices had fewer than 125 beneficiaries attributed to them, suggesting that applying a minimum of 125 beneficiaries would be overly restrictive for the intervention group and the comparison pool.

[44] We adjudicated all practices with missing specialty information and practices that were considered specialty practices to confirm their classification in the OneKey™ data. Primary care practices include practices with the following specialty categories: family practice, general practice, geriatric medicine, internal medicine, pediatric medicine, preventive medicine, and primary care.

Exhibit E.8. Flowchart of restrictions to non-MDPCP practices



Note: The percentages in the left column (green) refer to the percentage remaining of the total initial practice count, and the percentages in the right column (cream) refer to the percentage removed from the prior step because of the restriction.

MDPCP = Maryland Primary Care Program; NPI = National Provider Identifier; TIN = Tax Identification Number; VHA = Veterans Health Administration.

E.1.5b. Identifying variables to match on and setting criteria for what counts as sufficient balance

In close collaboration with CMS, we set priorities for matching variables and variables to check balance on to make the matching process feasible and on target (summarized in Exhibit E.10). So that the matched comparison group would estimate MDPCP's counterfactual, we set out to select a comparison group that had good balance on the following (even if we did not directly match on some, like outcomes):

- Beneficiary characteristics on aggregate, such as percentage of beneficiaries assigned to the practice who are female or mean HCC score
- Three priority non-outcome practice-level characteristics that were significantly imbalanced in the pre-matched sample: whether the practice belonged to a health system, Shared Savings Program (SSP) participation, and the number of affiliated providers (and thus, beneficiaries)
- Other non-outcome practice-level characteristics, such as whether the practice is a multi-specialty group practice or if the practice is primary care specialty
- Practice-level baseline (2017–2018) outcomes and trends (variables to check balance on)

- Practice-level baseline outcomes and trends for high-risk, high-needs beneficiaries subgroup, who are identified based on the Health Equity Advancement Resource and Transformation (HEART) payment eligibility criteria[45] (variables to check balance on)

In addition, we identified what we would count as sufficient balance for each of the matching variables. The method we used to reweight comparison practices allowed us to set balance standards for each individual variable. For beneficiary and most practice characteristics, we preferred to have standardized differences between the intervention and control groups of less than 0.15 but considered 0.25 to be sufficient. Whether the practice belonged to a health system, SSP participation, and the number of affiliated providers were identified as key variables that could have a strong influence on outcomes. We considered a range of balance targets for these variables because they were significantly imbalanced prior to matching to assess the trade-offs between achieving tighter balance versus increasing the sample size of the comparison group. The practice-level baseline outcomes and trends were not explicitly used in the reweighting algorithm because these were already well-balanced and because of regression to the mean concerns (Daw and Hatfield 2018). Still, for these variables, we wanted to see standardized differences of no larger than 0.25 and preferably less than 0.15 to strengthen the credibility of the parallel-trends assumption. We also wanted to make sure the balance for the baseline outcomes and trends was good within the subgroup of high-risk, high-needs beneficiaries to make future estimates about this subgroup more credible. However, because of the limited size of the comparison practice pool and issues with balance on other key matching variables, we set a looser standard for these of 0.25 standardized differences or smaller (Exhibit E.9).

Exhibit E.9. Baseline measures for selecting practices into the matched comparison group

Domain and measure	Data source	Included in matching[a]	SD required[b]
Average HCC score among beneficiaries assigned to the practice in 2018	Medicare claims	X	0.15
Average number of physical health chronic conditions 2018	Medicare claims	X	0.15
Percentage of beneficiaries who are dually eligible for Medicare and Medicaid in 2018	Medicare enrollment	X	0.15
Percentage of beneficiaries who are non-Hispanic Black in 2018	Medicare enrollment	X	0.15
Percentage of beneficiaries who live in a rural zip code in 2018	Medicare enrollment	X	0.15
Percent of beneficiaries assigned to the practice that are female in 2018	Medicare enrollment	X	0.15
Percentage of beneficiaries with original reason for Medicare entitlement as disability in 2018	Medicare enrollment	X	0.15
Average age of beneficiaries assigned to the practice as of January 2018	Medicare enrollment		<0.25

[45] High-risk, high-needs beneficiaries are those who meet the two criteria for HEART payments: (1) high risk (defined as a HCC score at or above the 75th percentile of all Maryland FFS beneficiaries) and (2) high needs (defined as living in an area in the highest quintile of the area deprivation index among all MDPCP beneficiaries).

Domain and measure	Data source	Included in matching[a]	SD required[b]
Percent of beneficiaries assigned to the practice that are Hispanic in 2018	Medicare enrollment		<0.25
Percent of beneficiaries assigned to the practice that are Asian/Pacific Islander in 2018	Medicare enrollment		<0.25
Percentage of beneficiaries assigned to the practice with original reason for Medicare entitlement as aged in 2018	Medicare enrollment		<0.25
Social Vulnerability Index		X	0.15
Count of NPIs at the practice	OneKey™, NPPES	X	0.30
Practice Health System indicator	OneKey™	X	0.30
Percentage of NPIs at the practice that are PCPs	OneKey™	X	0.30
SSP participation	MDM	X	0.30
Practice is multi-specialty group practice	OneKey™	X	0.15
Practice is primary care specialty	OneKey™	X	0.15
Practice located in Baltimore region Maryland	OneKey™	X	0.15
Practice located in Capital Region Maryland	OneKey™	X	0.15
Practice located in Central Maryland	OneKey™	X	0.15
Practice located in Eastern Shore Maryland	OneKey™	X	0.15
Practice located in Southern Maryland	OneKey™	X	0.15
Practice located in Western Maryland	OneKey™	X	0.15
Assigned beneficiary count in 2018 - unweighted	OneKey™, MD-PPAS, Medicare claims		<0.25
Practice does not have a corporate owner	OneKey™		<0.25
All ED visits 2017	Medicare claims		<0.25
All ED visits 2018	Medicare claims		<0.25
All ED visits slope (2018–2017)	Medicare claims		<0.25
All-cause hospital admissions 2017	Medicare claims		<0.25
All-cause hospital admissions 2018	Medicare claims		<0.25
All-cause hospital admissions slope (2018–2017)	Medicare claims		<0.25
Follow up after acute exacerbation of chronic conditions 2017	Medicare claims		<0.25
Follow up after acute exacerbation of chronic conditions 2018	Medicare claims		<0.25
Follow up slope (2018–2017)	Medicare claims		<0.25
Non-hospital spending 2017	Medicare claims		<0.25
Non-hospital spending 2018	Medicare claims		<0.25
Non-hospital spending slope (2017–2018)	Medicare claims		<0.25
PQI admissions 2017	Medicare claims		<0.25
PQI admissions 2018	Medicare claims		<0.25
PQI slope (2018–2017)	Medicare claims		<0.25
Preventable ED visits 2017	Medicare claims		<0.25

Domain and measure	Data source	Included in matching[a]	SD required[b]
Preventable ED visits 2018	Medicare claims		<0.25
Preventable ED slope (2018–2017)	Medicare claims		<0.25
Standardized hospital spending 2017	Medicare claims		<0.25
Standardized hospital spending 2018	Medicare claims		<0.25
Standardized hospital spending slope (2018–2017)	Medicare claims		<0.25
Total spending (non-hospital spending + standardized hospital spending) 2017	Medicare claims		<0.25
Total spending (non-hospital spending + standardized hospital spending) 2018	Medicare claims		<0.25
Total spending (non-hospital spending + standardized hospital spending) slope (2017–2018)	Medicare claims		<0.25
Unplanned readmissions 2017	Medicare claims		<0.25
Unplanned readmissions 2018	Medicare claims		<0.25
Unplanned readmissions slope (2018–2017)	Medicare claims		<0.25

Note: We conducted matching at the practice level. When applicable, we aggregated data to the practice level before analyzing or matching.

[a] Indicates whether the variable was included in the reweighting algorithm. Many variables listed in the table were not included the reweighting algorithm but were important to check for balance.

[b] The column "SD" refers to the maximum standardized differences we allow between the MDPCP and the comparison group. In our reweighting algorithm, we can set tolerances for individual variables to be more (lower SD) or less (higher SD) similar between MDPCP and the control group (see the section on reweighting method below for more details). We aimed for a standard of 0.25 SDs where possible, but some variables were too difficult to match on (required large tradeoffs in balance elsewhere or size of the comparison group) and thus were allowed to be more dissimilar on standardized differences (e.g., whether the practice belonged to a health system). For variables not included in the matching, the SD required represents our target for acceptable balance.

ED = emergency department; HCC = Hierarchical Condition Category; MDM = Medical Decision Making provider extract; MD-PPAS = Medicare Data on Provider and Practice Specialty; NPI = National Provider Identifier; NPPES = National Provider Plan Enumeration System; PCP = primary care provider; PQI = Prevention Quality Indicator; SSP = Shared Savings Program.

E.1.5c. Reweighting comparison practices to create the matched comparison group

To select our comparison group, we used the same methods for reweighting as was used to develop the comparison group for statewide effects (Appendix B, Section B.1.3). Like for the statewide impacts comparison group, stable balancing weights was selected over traditional matching techniques for MDPCP because it allows for tailored balance criteria for each matching variable. This tailoring enabled us to identify and make precise tradeoffs between balance on select variables versus the size and distribution of the comparison group, including those that were difficult to match because they were not well balanced at baseline, such as health system.

Using an optimization-based approach, theoretically, any number of criteria can be set as constraints. As we add constraints (or tighten or require greater similarity between treated and comparison groups), however, the optimization problem becomes more difficult. The tradeoff to higher degrees of similarity across many different criteria is often the size of the comparison group represented. In other words, the algorithm will start to drop (that is, assign zero weight to) units that are too different from its target when there are no other options.

E.1.5d. Assessing the quality of the matched comparison group in terms of size, geographic spread, health system composition, balance, and statistical power

In selecting the comparison group, we aimed for a group that:

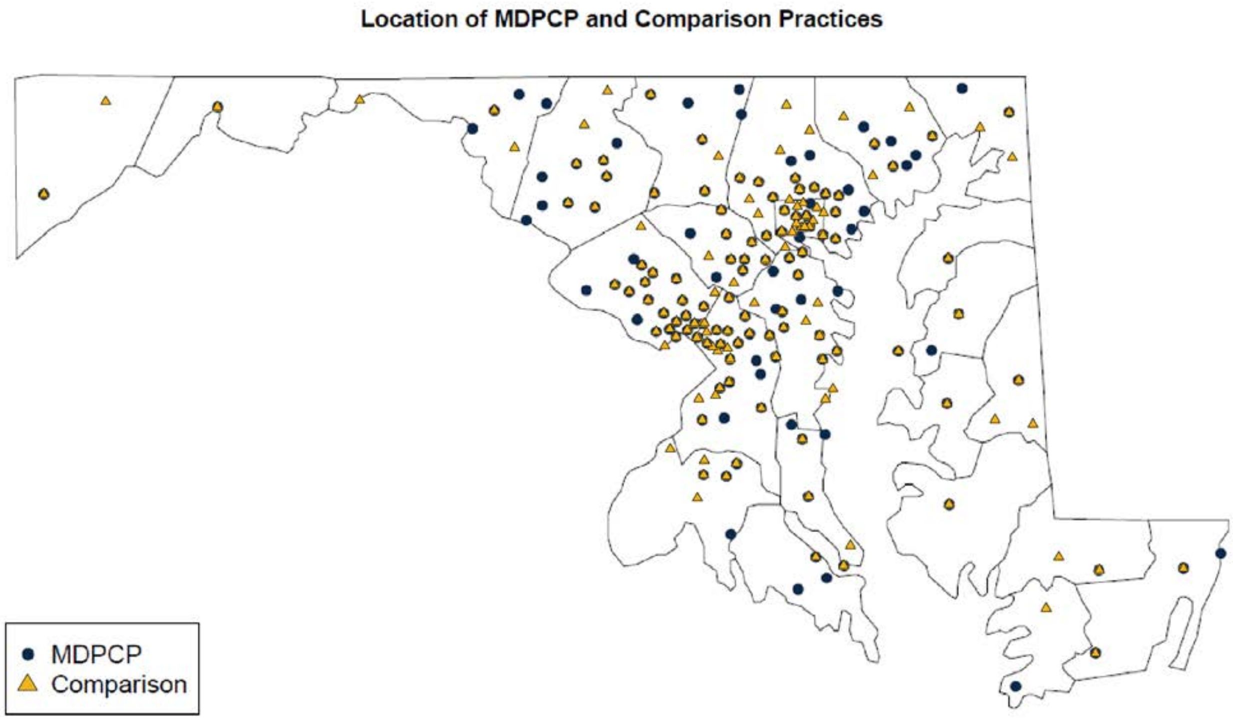
1. Had sufficient balance on all variables listed as priorities for matching.
2. Was large enough to support statistical inference and spread across Maryland geographically with a reasonable mix of health systems. These features help improve statistical power and avoid the possibility that idiosyncratic health shocks in any one area or health system would drive the results.

Conditions (1) and (2) generally trade off with one another—with more precise balance coming at the expense of a smaller and less dispersed comparison group (and worse statistical power). We explored several alternative comparison groups with CMS and selected the one that represented the best tradeoffs across these dimensions.

Size and geographic distribution of the comparison group

The selected comparison group was sufficiently large to support inference and spread well across Maryland geographically. The selected comparison group achieved an effective comparison group sample size of 261 practices. This number is lower than the number of MDPCP practices (implying a less than 1:1 match) but still sufficient to estimate impacts (power is discussed further below). Exhibit E.10 is a map of the MDPCP and comparison practices. The selected comparisons cover most of Maryland and are distributed similarly as the MDPCP practices. The matched comparison group also captured a good mix of health systems. Of the health systems in the comparison group, about 70% of the weight concentrates in the top five largest weighted health systems, which includes Mercy Medical Center, the Johns Hopkins Health System, and the University of Maryland Medical System. These specific health systems also had practices participating in MDPCP. Including these practices in the comparison group risks some amount of spillover, but it also might make them strong comparisons. Overall, we believe the benefits of including these system practices, as well as the risks to our comparison group of not having at least moderate balance on percentage of practices in health systems, outweighs the risks from spillover; however, we conduct a sensitivity analysis that removes these practices (detailed in Section E.1.11).

Exhibit E.10. Location of MDPCP and comparison practices in Maryland



MDPCP = Maryland Primary Care Program.

Balance on key characteristics and outcomes

Overall, we achieved good beneficiary-level balance in our selected comparison group (Exhibit E.11). For most beneficiary and practice characteristics, we achieved tight balance (within 0.15 standardized differences). The three priority non-outcome practice-level characteristics that were significantly imbalanced in the pre-matched sample—whether the practice belonged to a health system, SSP participation, and number of providers—were moderately well balanced with standardized differences at or below 0.33 after matching. Without having directly matched on them, outcome levels and trends were well balanced with standardized differences within 0.15. For the high-risk, high-needs subgroup, for most outcomes, balance levels and trends were similar to the full group, with some variables having slightly worse (for example, non-hospital spending among high-risk, high-needs beneficiaries had a post-weighted standardized difference of 0.23) but still acceptable balance.

Exhibit E.11. Balance between MDTCOC and selected comparison group on key characteristics and outcomes

Variable description	MDPCP pre-weighted mean	Comparison pre-weighted mean	Difference pre-weighting	Standardized difference pre-weighting	Comparison post-weighted mean	Difference post-weighting	Standardized difference post-weighting
Average HCC score among beneficiaries assigned to the practice in 2018	1.17	1.18	-0.01	-0.03	1.17	0.01	0.04
Average number of physical health chronic conditions 2018	4.79	4.88	-0.09	-0.12	4.79	-0.01	-0.01
Percentage of beneficiaries who are dually eligible for Medicare and Medicaid in 2018	0.13	0.17	-0.04	-0.26	0.15	-0.02	-0.14
Percentage of beneficiaries who are non-Hispanic Black in 2018	0.20	0.22	-0.01	-0.06	0.21	-0.01	-0.05
Percentage of beneficiaries who live in a rural zip code in 2018	16.77	17.32	-0.56	-0.03	18.31	-1.55	-0.08
Percent of beneficiaries assigned to the practice that are female in 2018	0.59	0.59	0.01	0.09	0.59	0.00	0.05
Percentage of beneficiaries with original reason for Medicare entitlement as disability in 2018	0.17	0.17	0.00	-0.04	0.17	0.00	-0.02
Count of NPIs at the practice	9.13	3.59	5.54	0.70	6.51	2.63	0.33
Practice Health System indicator	0.48	0.11	0.37	0.83	0.34	0.14	0.32
Percentage of NPIs at the practice that are PCPs	0.91	0.94	-0.03	-0.22	0.90	0.01	0.04
SSP participation	0.37	0.13	0.24	0.57	0.23	0.14	0.32
Social Vulnerability Index	0.36	0.38	-0.02	-0.19	0.38	-0.02	-0.15
Practice is multi-specialty group practice	0.13	0.06	0.08	0.26	0.13	0.00	0.00
Practice is primary care specialty	0.87	0.94	-0.08	-0.26	0.87	0.00	0.00
Practice located in Baltimore region Maryland	0.09	0.11	-0.02	-0.08	0.12	-0.03	-0.10
Practice located in Capital Region Maryland	0.24	0.27	-0.03	-0.07	0.24	0.01	0.02
Practice located in Central Maryland	0.39	0.35	0.04	0.08	0.39	0.00	0.00
Practice located in Eastern Shore Maryland	0.10	0.11	-0.02	-0.06	0.11	-0.01	-0.03
Practice located in Southern Maryland	0.07	0.05	0.02	0.10	0.04	0.03	0.14
Practice located in Western Maryland	0.11	0.10	0.01	0.03	0.12	0.00	0.00

Variable description	MDPCP pre-weighted mean	Comparison pre-weighted mean	Difference pre-weighting	Standardized difference pre-weighting	Comparison post-weighted mean	Difference post-weighting	Standardized difference post-weighting
Average age of beneficiaries assigned to the practice as of January 2018	72.89	73.18	-0.29	-0.08	73.07	-0.18	-0.05
Assigned beneficiary count in 2018 - unweighted	1,281.81	510.52	771.29	0.87	654.17	627.64	0.71
Medicare Advantage penetration rate	0.14	0.15	-0.01	-0.11	0.15	-0.01	-0.13
Count of PCPs at the practice	7.59	3.03	4.56	0.82	5.06	2.53	0.46
Practice does not have a corporate owner	0.39	0.83	-0.44	-0.91	0.59	-0.20	-0.41
Percent of beneficiaries assigned to the practice that are Hispanic in 2018	0.02	0.03	-0.01	-0.14	0.02	0.00	-0.06
Percent of beneficiaries assigned to the practice that are Asian/Pacific Islander in 2018	0.03	0.06	-0.03	-0.25	0.05	-0.02	-0.16
Percentage of beneficiaries assigned to the practice with original reason for Medicare entitlement as aged in 2018	0.83	0.82	0.00	0.03	0.83	0.00	0.02
All-cause hospital admissions 2017	289.22	287.09	2.13	0.02	282.19	7.03	0.08
All-cause hospital admissions 2018	276.59	277.70	-1.11	-0.01	276.61	-0.02	0.00
All ED visits 2017	475.09	475.38	-0.29	0.00	473.44	1.65	0.01
All ED visits 2018	474.83	483.59	-8.75	-0.04	481.49	-6.66	-0.03
Preventable ED visits 2017	185.22	183.38	1.84	0.02	182.60	2.61	0.03
Preventable ED visits 2018	182.29	183.74	-1.46	-0.02	182.66	-0.38	0.00
PQI admissions 2017	51.68	52.12	-0.45	-0.02	50.51	1.16	0.04
PQI admissions 2018	47.77	48.01	-0.24	-0.01	48.05	-0.28	-0.01
Non-hospital spending 2017	6,074.31	6,018.29	56.02	0.04	5,860.80	213.52	0.15
Non-hospital spending 2018	6,397.53	6,288.07	109.47	0.07	6,167.89	229.65	0.15
Standardized hospital spending 2017	5,026.96	4,899.75	127.21	0.08	4,921.56	105.40	0.07
Standardized hospital spending 2018	5,020.84	4,914.13	106.71	0.07	4,974.23	46.61	0.03
Total spending (non-hospital spending + standardized hospital spending) 2017	11,101.28	10,918.04	183.24	0.07	10,782.36	318.92	0.12
Total spending (non-hospital spending + standardized hospital spending) 2018	11,418.38	11,202.20	216.18	0.09	11,142.12	276.25	0.11

Variable description	MDPCP pre-weighted mean	Comparison pre-weighted mean	Difference pre-weighting	Standardized difference pre-weighting	Comparison post-weighted mean	Difference post-weighting	Standardized difference post-weighting
Follow up after acute exacerbation of chronic conditions 2017	76.28	74.91	1.37	0.14	75.81	0.47	0.05
Follow up after acute exacerbation of chronic conditions 2018	75.97	74.03	1.94	0.19	74.51	1.46	0.14
Unplanned readmissions 2017	15.40	15.36	0.04	0.01	15.09	0.31	0.06
Unplanned readmissions 2018	15.22	15.52	-0.30	-0.06	15.65	-0.43	-0.09
All-cause hospital admissions slope (2018–2017)	-12.63	-9.39	-3.24	-0.06	-5.57	-7.06	-0.14
All ED visits slope (2018–2017)	-0.25	8.21	-8.46	-0.10	8.06	-8.31	-0.10
Preventable ED slope (2018–2017)	-2.93	0.37	-3.30	-0.08	0.06	-2.99	-0.07
PQI admissions slope (2018–2017)	-3.91	-4.11	0.21	0.01	-2.47	-1.44	-0.07
Non-hospital spending slope (2017–2018)	323.22	269.78	53.44	0.08	307.09	16.13	0.02
Standardized hospital spending slope (2018–2017)	-6.12	14.38	-20.50	-0.02	52.67	-58.79	-0.06
Total spending (non-hospital spending + standardized hospital spending) slope (2017–2018)	317.10	284.16	32.94	0.05	359.76	-42.66	-0.08
Follow up slope (2018–2017)	-0.31	-0.88	0.57	0.05	-1.30	0.99	0.08
Unplanned readmissions slope (2018–2017)	-0.18	0.16	-0.34	-0.06	0.56	-0.74	-0.12

Note: The pre-weighted means are the raw practice-level means (weighted only for FFS beneficiary count). Post-weighted means are weighted by the final matching weights.

ED = emergency department; FFS = fee for service; HCC = Hierarchical Condition Category; NPI = National Provider Identifier; PCP = primary care provider; PQI = Prevention Quality Indicator; SSP = Shared Savings Program.

E.1.6. Power

We determined the minimum detectable effect for the MDPCP analysis—that is, the smallest effect, in percentage terms, that we expect to be able to detect under a range of assumptions about the relative size of the comparison group and the intracluster correlation coefficient (ICC).[46] We focused on all-cause admissions for these calculations. Other assumptions include the following: 361 MDPCP practices, an average of 471 assigned beneficiaries per practice, a coefficient of variation of 3.0 based on values reported from broad literature review in Peikes et al. (2011), and an estimated R-squared value of 0.13. These assumptions are based on prior work for impacts and power calculations of primary care interventions, such as Comprehensive Primary Care Plus and Primary Care First.

At an ICC of 0.002 (our best estimate based on findings from Comprehensive Primary Care Plus), we will have 80% power for a two-sided test at the 10% level to detect effects on all-cause admissions of about 3.5% with a sample size of about 250 comparison practices (our matched comparison group has an effective sample size of 261). An alternative matching option was considered that produced a larger effective sample size; however, this came at the cost of significantly worse balance on the tough-to-match key practice characteristics. Power for a high-risk, high-needs beneficiary sub-analysis is similar relative to the power for the overall MDPCP impacts analysis, but it would require larger effects in that group. We assumed the following: 358 MDPCP practices (three practices have no high-risk, high needs beneficiaries); a mean of 31 high-risk, high-needs beneficiaries per practice; a coefficient of variation of 3.0; and an R-squared of 0.13. At an ICC of 0.002, the minimum detectable effect is about 4.5% with a comparison group sample size of about 250 practices.

[46] The ICC measures how much the clusters (in our case, practices) can explain the variation in outcomes across people in the study population. Higher ICCs mean that outcomes among people within a cluster tend to be more similar than outcomes among people across clusters.

E.1.7. Regression specifications

E.1.7a. Regression specifications and statistical testing for beneficiary-year and episode-year Medicare FFS claims-based analyses

We used linear regression models to implement the difference-in-differences impact analyses. We measured impacts separately for each year and for the combined years of 2019–2020 and 2021–2022 as well as a combined four-year effect (2019–2022)[47]. The Chapter 5 findings included two units of analysis: (1) analyses of observations for each Medicare FFS beneficiary in MDPCP and the matched comparison regions for each year, and (2) analyses of episode outcomes with observations for each episode for each year (episode-year analyses). The beneficiary-year and episode-year models accounted for the clustering of beneficiaries within practices through cluster-robust standard errors, controlled for time-invariant effects of unobserved confounders and common shocks through the use of fixed effects, and they included baseline and time-varying covariates as independent variables.

Impact estimates

The difference-in-differences regression models for the beneficiary-year analyses with claims-based outcome measures used Medicare FFS data with one observation per beneficiary for each year (2017 to 2022). The regression models for the episode-year analysis took the same form, but with the unit of analysis as the episode rather than the beneficiary. The regression model to estimate the yearly impact for beneficiary- and episode-level estimates took the following form:

$$(1) \qquad y_{it} = \sum\nolimits_{\tau=2019}^{2022} T_\tau M_r \delta_{t,\tau} + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

In this model, y_{it} represents the outcome for beneficiary *i* (or episode *i*) in year *t* in practice r, τ indexes years (with $\tau = 2019$ corresponding to the first year),[48] M_r equals 1 for MDPCP beneficiaries (or episodes) and 0 for beneficiaries (or episodes) from the comparison regions, and T_τ is a dummy variable that equals 1 for observations in year τ and equals 0 otherwise. X_{it} is a set of independent covariates whose relationship with the outcome we allow to change with time using an interaction term. γ_t represents a set of year fixed effects and μ_r represents a set of practice-level fixed effects for beneficiary-year and episode-year outcomes.

[47] Standardized spending data is only available through 2021, so the combined years include 2019–2020 and the individual 2021 year.

[48] All time trends are relative to the last year of the baseline period (2018), which is the reference year in the regression models.

Beneficiaries in MDPCP generally receive a weight of 1 in the regression models. But in cases in which a beneficiary is unobservable (that is, not alive and enrolled in Medicare Parts A and B with Medicare as their primary payer) the whole year, we annualized their beneficiary-year outcomes and constructed observability weights that reflect the amount of time that the beneficiary is observable in the year. For the comparison group beneficiaries, we applied the matching weights (detailed in Section E.1.4) to account for the practice-level reweighting along with the observability weights; the two weights were multiplied together to produce a final, beneficiary-level weight. For episode analyses, we applied the matching weights to comparison group beneficiaries, and MDPCP beneficiaries received a weight of 1 because episode analyses were not annualized.

The impact estimates are the δs—the change in mean outcomes in the intervention group each year after accounting for the changes in the comparison group in the respective year (the γ_ts). Separate estimates for each year (that is, one δ per year) allowed for nonlinearity in the effects (for example, effects might not occur immediately or could level off or decline over time). In addition to the yearly impact estimates, we also estimated the combined effect during 2019–2020 and 2021–2022 and across the first four years of MDPCP (2019–2022).

The regression model to estimate the combined 2019–2022 impact estimates took the following form:

$$(2) \qquad y_{it} = T_{2019-2022} M_r \delta_\gamma + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

In this model, y_{it} represents the outcome for beneficiary *i* in year *t* in practice *r*, M_r equals 1 for MDPCP beneficiaries and 0 for beneficiaries from the comparison regions, $T_{2019-2022}$ is a dummy variable that equals 1 for observations in years 2019 to 2022. X_{it} is a set of independent covariates whose relationship with the outcome we allow to change with time using an interaction term. γ_t represents a set of year fixed effects and μ_r represents a set of practice-level fixed effects for beneficiary-year outcomes and hospital fixed effects for episode-year outcomes. δ_γ represents the impact estimates across the first four years of MDPCP.

The regression model to estimate the combined 2019–2020 and 2021–2022 impact estimates took the following form:

$$(3) \qquad y_{it} = T_{2019-2020} M_r \delta_\gamma + T_{2021-2022} M_r \delta_\gamma + X_{it}\beta + \gamma_t + \mu_r + \varepsilon_{it}$$

In this model, y_{it} represents the outcome for beneficiary *i* in year *t* in practice *r*, M_r equals 1 for MDPCP beneficiaries and 0 for beneficiaries from the comparison regions, $T_{2019-2020}$ is a dummy variable that equals 1 for observations in years 2019 to 2020, and $T_{2021-2022}$ is a dummy variable that equals 1 for observations in years 2021 to 2022. X_{it} is a set of independent covariates whose relationship with the outcome we allow to change with time using an interaction term. γ_t represents a set of year fixed effects and μ_r represents a set of practice-level fixed effects for beneficiary-year outcomes and hospital fixed effects for episode-year outcomes. δ_{γ} represents the impact estimates across the first four years of MDPCP.

E.1.7b.Regression-adjusted means and percentage impact

We calculated regression-adjusted means and percentage impact for each of our estimated outcomes. The percentage impact is calculated using the same approach as the main impact analysis (see Appendix B, Section B.1.7b for details). In all periods, including baseline (2017–2018) and the intervention period (2019–2022), the regression-adjusted mean for Maryland is the mean of the outcome in MDPCP during that period or year (weighted for observability in claims-based beneficiary-year analyses). For the comparison group, in the baseline period, we calculated the regression-adjusted mean as the mean of the outcome in the comparison group weighted by the practice matching weights (times observability in claims-based beneficiary-year analyses). In all post-baseline years (2019–2022), we calculated the regression-adjusted mean in the comparison group as the MDPCP mean in that period or year minus the difference-in-differences impact estimate associated with that period or year, minus the difference between MDPCP and the comparison group in the baseline period.

E.1.8. Additional measures, definitions, and analytic file construction

Like for the statewide impacts analysis, we fit regression models at the beneficiary-year or discharge-year level, as relevant, for the claims-based outcomes. The beneficiary-year-level analytic file contains one observation per beneficiary per year for all beneficiaries enrolled in Medicare in that year who are assigned to either the 2019 MDPCP practices or the matched comparison group. For the analyses of claims-based measures, we then limited the file to those who were observable for at least one month in Medicare FFS claims data during the year (that is, they were alive, enrolled in Medicare Parts A and B FFS, and had Medicare as primary payer). Beneficiaries can be in the file in all years of our analytic period or only one or a limited number of years, depending on their observability status. We use the beneficiary-level and discharge-level analytic files from the main analysis described in Appendix B but fix the covariates for each beneficiary at the point of assign (the first observation in the baseline period [2017–2018] and the first observation in the intervention period [2019–2022]). We also use the same regression covariates as the main beneficiary-level and discharge-level analyses (see Exhibit B.12 in Appendix B.1.6) and include additional practice-level covariates (Exhibit E.12).

Exhibit E.12. Practice-level covariates for beneficiary- and discharge-level regression models

Covariate	Definition	Data source
Practice NPI count	The count of NPIs participating in MDPCP at the practice	OneKey™
Practice percent PCP	The percentage of NPIs at the practice that are primary care providers	OneKey™ and NPPES
Practice health system	Indicator for whether the practice is affiliated with a health system	OneKey™
Practice independent	Indicator for whether the practice has a corporate owner	OneKey™
Practice SSP participation	Indicator for whether the practice participates in a Shared Savings Program	MDM
Practice rural percent	Percentage of beneficiaries attributed to the practice that live in a rural area	Medicare enrollment and claims data

MDM = Medical Decision Making provider extract; MDPCP = Maryland Primary Care Program; NPI = National Provider Identifier; NPPES = National Plan and Provider Enumeration System; PCP = primary care provider; SSP = Shared Savings Program.

Number of non-emergent or primary care treatable outpatient emergency department visits and observation stays (number of visits per beneficiary per year)

One outcome measure that is unique to the MDPCP impacts analysis is the number of non-emergent or primary care treatable emergency department visits and observation stays. This measure is the annualized number of outpatient emergency department visits and observation stays that a beneficiary had in a year classified as non-emergent or primary care treatable by the algorithm published by New York University (Billings et al. 2000a, 2000b). To construct this measure, we applied the patched version of the New York University software algorithm to outpatient emergency department visits and observation stays (Johnston et al. 2017). For each visit, the algorithm assigns a specific percentage of the visit into the following categories:

1. Non-emergent
2. Emergent/primary care treatable
3. Emergent, emergency department care needed, preventable/avoidable
4. Emergent, emergency department care needed, not preventable/avoidable[49]

For example, a given visit might be assigned 10% to category 1, 50% to category 2, 30% to category 3, and 10% to category 4. For each beneficiary, we calculated the sum of the percentages for categories 1 and 2 across all the emergency department visits and observation stays during the year.

E.1.9. Unadjusted means by year

To help interpret what drives the difference-in-differences impact estimates, we include the size of the intervention and comparison groups over time, including the number of beneficiaries in the high-risk, high-needs subgroup (Exhibit E.13) and the trends in unadjusted (but comparison group weighted) means for outcome measures since 2017 (Exhibits E.14 and E.15) for these populations. For the means figures, MDPCP beneficiaries are weighted by their observability in the year, and

[49] The algorithm first identifies visits that are for injuries or are related to mental health, drugs, or alcohol. Johnston et al. (2017) found that about 26% of all emergency department visits are for these conditions. Further, another 8% of emergency department visits cannot be classified because their diagnosis codes do not map to one of the four categories listed above. Therefore, nationally, the New York University algorithm does not assign one of the four categories above to roughly 34% (26% + 8%) of all emergency department visits.

beneficiaries in the comparison group are weighted by their observability and matching weights. Episode-level outcomes (30-day unplanned readmission and timely follow-up after acute exacerbation of a chronic condition), which are not annualized, receive a weight of 1 for MDPCP episodes, while episodes in the comparison group are weighted by their matching weights.

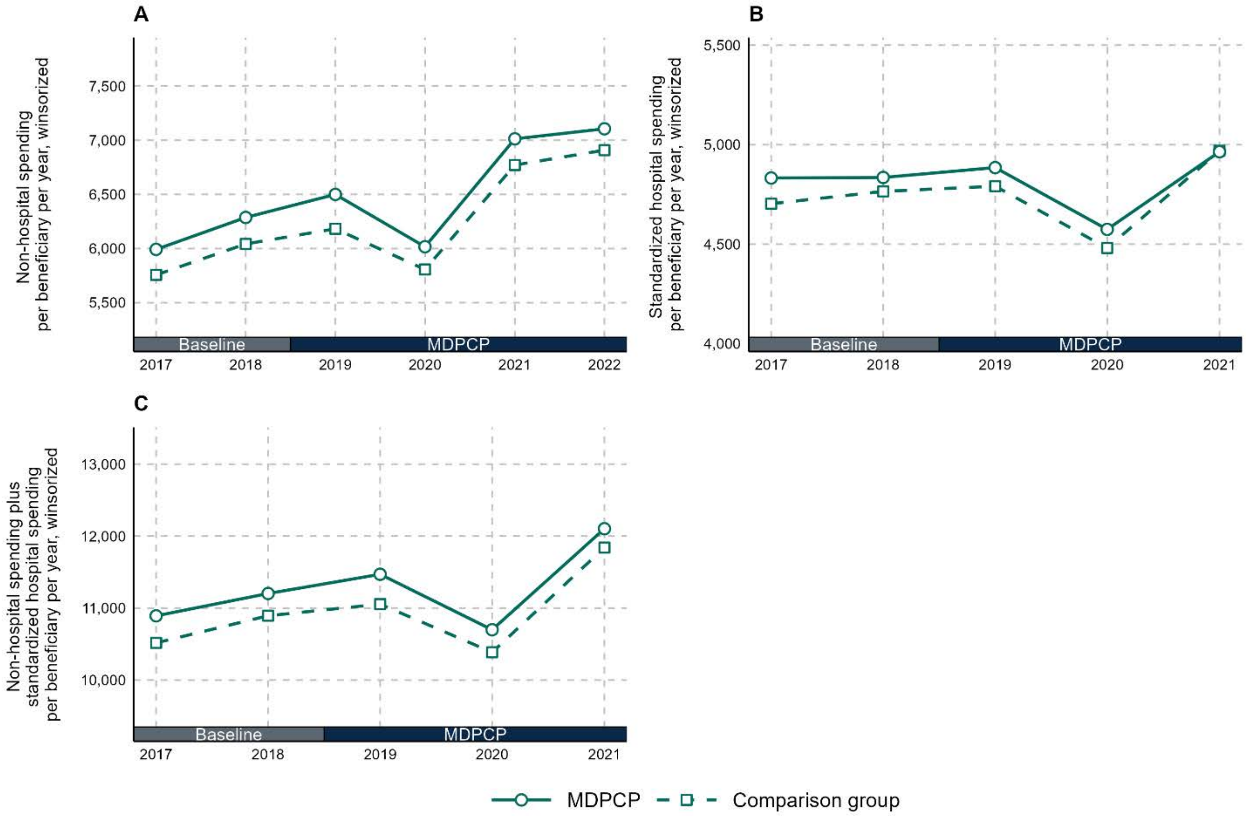
For the beneficiary analyses, MDPCP counts steadily increase over time, while the matched comparison stays approximately the same. The episode counts have decreased slightly in both MDPCP and the comparison group since 2019, reflecting decreases in utilization associated with COVID-19.

Exhibit E.13. Size of the MDPCP and comparison groups over time (weighted)

Year	Weighted MDPCP count	Weighted comparison count
Beneficiary-level analysis counts (number of Medicare FFS beneficiaries)		
2017	223,254	114,634
2018	249,412	124,177
2019	263,769	115,447
2020	292,507	122,743
2021	292,329	120,629
2022	292,563	117,475
Episode analysis: 30-day post-discharge unplanned readmission index admission counts (number of index admissions)		
2017	60,140	29,826
2018	64,121	31,554
2019	64,833	27,921
2020	60,011	24,477
2021	61,625	25,567
2022	61,508	25,305
Episode analysis: Timely follow-up after acute exacerbation of chronic conditions denominator counts (number of eligible index admissions and ED visits)		
2017	23,685	12,008
2018	25,648	12,847
2019	26,764	11,791
2020	22,568	9,130
2021	22,270	9,188
2022	21,504	8,720
Subgroup analysis: High-risk, high-needs beneficiary-level analyses		
2017	13,143	7,641
2018	14,781	8,106
2019	16,345	7,377
2020	18,426	8,133
2021	18,835	8,123
2022	18,734	7,820

ED = emergency department; FFS = fee for service; MDPCP = Maryland Primary Care Program.

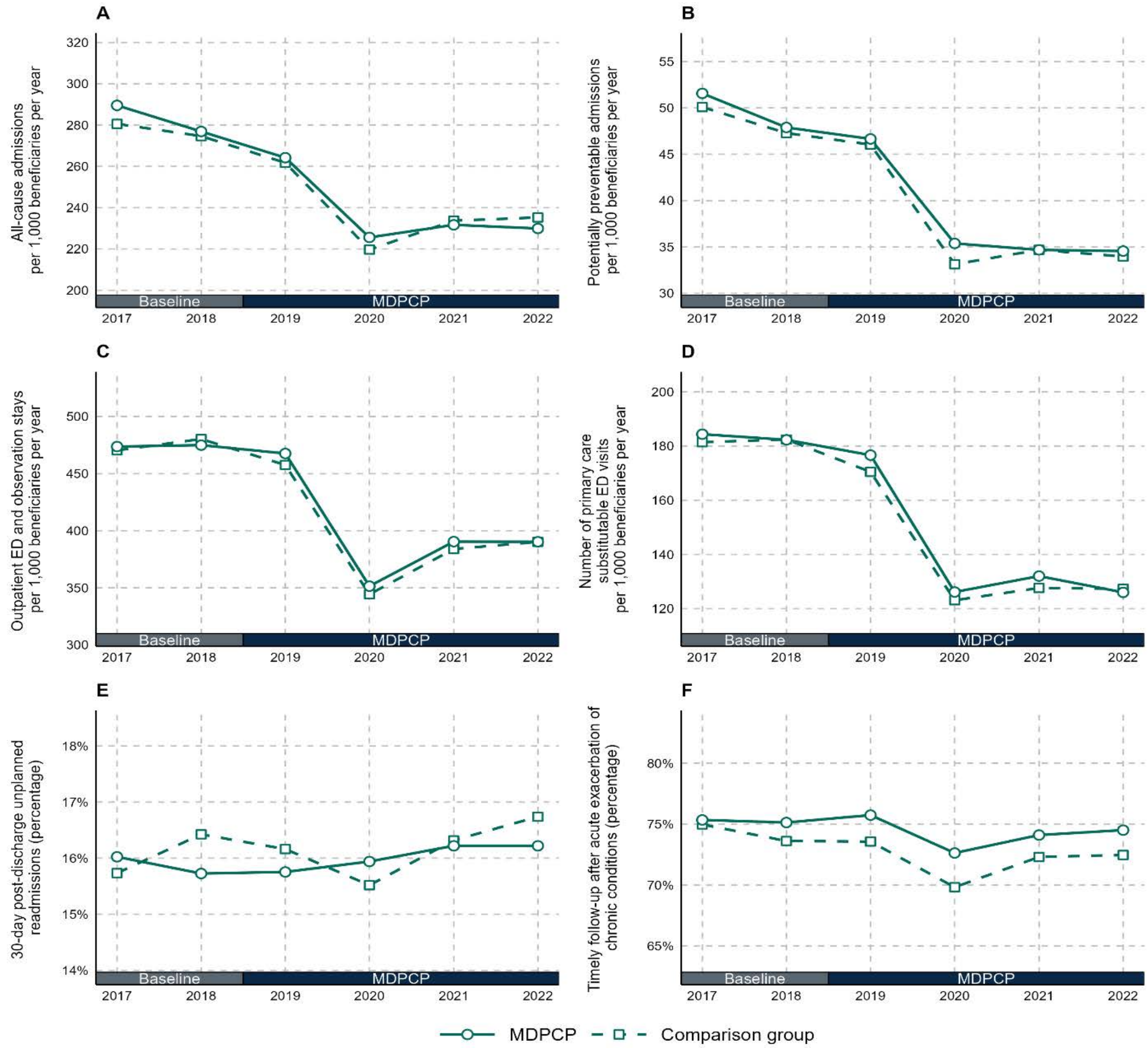
Exhibit E.14. Unadjusted spending per beneficiary per year after matching



Note: MDPCP mean is weighted for observability in Medicare FFS claims. Comparison group mean is weighted for matching and observability.

FFS = fee for service; MDPCP = Maryland Primary Care Program.

Exhibit E.15. Unadjusted utilization and quality outcomes after matching



Note: MDPCP mean is weighted for observability (except for 30-day unplanned readmissions and follow-up after acute exacerbation which are episode level). Comparison group mean is weighted for matching and observability.

ED = emergency department; MDPCP = Maryland Primary Care Program.

E.1.10. Tables of impact estimates and regression adjusted means, by year

In this section (Exhibit E.16), we present regression-adjusted means as well as impact estimates of the MDPCP by combined years (2019–2020, 2021–2022, 2019–2022). See Appendix B, Section B.1.8 for details on how the regression-adjusted means are calculated.

Exhibit E.16. MDPCP impact estimates

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
All-cause admissions (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	283	277	6		
MDPCP					
Combined (2019–2020)	244	243	1	-5 (-11; 1)	-2.0%
Combined (2021–2022)	231	232	-1	-7 (-14; 0)	-2.9%
Combined (2019–2022)	237	237	0	-6 (-12; 0)	-2.5%
PQI admissions (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	49.6	48.6	1.0		
MDPCP					
Combined (2019–2020)	40.7	40.4	0.3	-0.7 (-2.6; 1.2)	-1.7%
Combined (2021–2022)	34.6	34.6	0.0	-1 (-3.3; 1.3)	-2.8%
Combined (2019–2022)	37.6	37.4	0.2	-0.8 (-2.8; 1.1)	-2.1%
All-cause ED visits and observation stays (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	474	476	-2		
MDPCP					
Combined (2019–2020)	407	408	-1	1 (-10; 13)	0.2%
Combined (2021–2022)	390	398	-8	-6 (-18; 6)	-1.5%
Combined (2019–2022)	398	402	-4	-2 (-14; 9)	-0.5%
Non-emergent or primary care treatable ED visits (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	183	182	1		
MDPCP					
Combined (2019–2020)	150	149	1	0 (-5; 5)	-0.0%
Combined (2021–2022)	129	133	-4	-5 (-10; 0)	-3.7%
Combined (2019–2022)	139	140	-1	-2 (-7; 2)	-1.4%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Standardized hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$4,834	$4,735	$99		
MDPCP					
Combined (2019–2020)	$4,722	$4,679	$43	-56 (-155; 43)	-1.2%
Combined (2021)	$4,964	$5,003	-$39	-138** (-248; -29)	-2.7%
Combined (2019–2022)	$4,805	$4,790	$15	-84 (-171; 3)	-1.7%
Non-hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$6,147	$5,905	$242		
MDPCP					
Combined (2019–2020)	$6,245	$5,974	$271	29 (-72; 130)	0.5%
Combined (2021–2022)	$7,059	$6,736	$323	81 (-36; 197)	1.2%
Combined (2019–2022)	$6,663	$6,366	$297	55 (-47; 157)	0.8%
Non-hospital spending and standardized hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$11,056	$10,713	$343		
MDPCP					
Combined (2019–2020)	$11,064	$10,737	$327	-16 (-195; 163)	-0.1%
Combined (2021)	$12,103	$11,784	$319	-24 (-235; 186)	-0.2%
Combined (2019–2022)	$11,422	$11,098	$324	-19 (-191; 154)	-0.2%
30-day post-discharge unplanned readmission (percentage)					
Baseline period (2017–2018)	15.9%	16.1%	-0.2%		
MDPCP					
Combined (2019–2020)	15.8%	15.9%	-0.1%	0.1 (-0.5; 0.8)	0.6%
Combined (2021–2022)	16.2%	16.5%	-0.3%	-0.1 (-0.8; 0.5)	-0.6%
Combined (2019–2022)	16.0%	16.2%	-0.2%	0 (-0.6; 0.6)	-0.0%
Timely follow-up after acute exacerbation of a chronic condition (percentage)					
Baseline period (2017–2018)	75.2%	74.3%	0.9%		
MDPCP					
Combined (2019–2020)	74.3%	71.6%	2.7%	1.8** (0.6; 3)	2.5%
Combined (2021–2022)	74.3%	72.4%	1.9%	1 (-0.4; 2.3)	1.4%
Combined (2019–2022)	74.3%	72.0%	2.3%	1.4** (0.3; 2.6)	1.9%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year for MDPCP minus the difference-in-differences impact estimate for the year.

[b] Spending is winsorized to the 99.5th percentile.

CI = confidence interval; ED = emergency department; MDPCP = Maryland Primary Care Program; PQI = Prevention Quality Indicator.

E.1.11. Health system spillover sensitivity analysis

Prior to matching, MDPCP practices were far more likely to be a part of a health system than the pool of comparison practices (48% versus 11%). Health system affiliation was a high priority matching variable for the analysis, because whether a practice is part of a health system could influence outcome trend for their Medicare patients, independent of MDPCP participation. If we did not match on health system, we could conflate impacts of MDPCP with those of being part of a health system After matching, the balance on health system improved, although a moderate imbalance remained (48% in MDPDP versus 34% in the comparison group). To achieve this balance, several comparison group practices were affiliated with systems that had high uptake of MDPCP, defined as at least 75% of the system's MDPCP-eligible practices participating in MDPCP. In interviews with practices (see Appendix E.2), we found that some health systems have implemented some care changes to meet MDPCP requirements across all their practices, not only their practices in MDPCP. This could lead to some spillover of the MDPCP intervention to practices in our comparison group, making the measured effects smaller than true effects. In this sensitivity analysis, we remove practices from the comparison pool if more than 75% of the practices we identified in the system were participating in MDPCP. Comparison group practices in Johns Hopkins Health System, University of Maryland Medical System, GBMC Healthcare Inc., Medstar Health, and Atlantic General Health System are removed in the sensitivity analysis (N = 16 practices, 9.7% of weighted beneficiaries in the comparison group), resulting in the percent of beneficiaries in a health system practice shifting from 34% to 24%.

The full results for all outcomes where we remove the spillover practices are in Exhibit E.17. We lead with the main results (which include these practices at high-risk of spillover in the comparison group) because we pre-specified this design to limit the confounding influence of health systems on outcomes. Overall, we found similar results across the first four years of the program between the main analysis and the health system spillover over sensitivity analysis, except for the hospital use outcomes (all-cause admissions, potentially preventable admissions, and standard hospital spending), which had larger, statistically significant reductions in the sensitivity analysis. This suggests that there was possible spillover within health systems and that MDPCP impact estimates without any spillover may be slightly more favorable than what we observed in our primary analyses.

Exhibit E.17. MDPCP impact estimates after removing practices with the potential for health system spillover

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
All-cause admissions (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	283	274	9		
MDPCP					
Combined (2019–2020)	244	242	2	-7* (-13; -1)	-2.8%
Combined (2021–2022)	231	230	1	-8** (-15; -2)	-3.3%
Combined (2019–2022)	237	236	1	-8** (-13; -2)	-3.3%
PQI admissions (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	49.6	47.3	2.3		
MDPCP					
Combined (2019–2020)	40.7	40.3	0.4	-1.9 (-3.8; 0.1)	-4.5%
Combined (2021–2022)	34.6	34.4	0.2	-2.1 (-4.4; 0.1)	-5.7%
Combined (2019–2022)	37.6	37.3	0.3	-2* (-3.9; -0.1)	-5.1%
All-cause ED visits and observation stays (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	474	470	4		
MDPCP					
Combined (2019–2020)	407	405	2	-2 (-14; 10)	-0.5%
Combined (2021–2022)	390	395	-5	-9 (-21; 3)	-2.3%
Combined (2019–2022)	398	400	-2	-6 (-17; 6)	-1.5%
Non-emergent or primary care treatable ED visits (number per 1,000 beneficiaries per year)					
Baseline period (2017–2018)	183	179	4		
MDPCP					
Combined (2019–2020)	150	147	3	-1 (-7; 4)	-0.7%
Combined (2021–2022)	129	130	-1	-5* (-10; 0)	-3.7%
Combined (2019–2022)	139	138	1	-3 (-8; 1)	-2.1%
Standardized hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$4,834	$4,656	$178		
MDPCP					
Combined (2019–2020)	$4,721	$4,637	$84	-94 (-190; 2)	-2.0%
Combined (2021)	$4,964	$4,964	$0	-178** (-284; -71)	-3.5%
Combined (2019–2022)	$4,805	$4,750	$55	-123** (-204; -42)	-2.5%

	Regression-adjusted mean			Difference-in-differences impact estimate, by year	
	Maryland	Comparison group	Difference	Estimate (90% CI)	% Impact[a]
Non-hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$6,147	$5,896	$251		
MDPCP					
Combined (2019–2020)	$6,245	$5,994	$251	0 (-107; 107)	-0.0%
Combined (2021–2022)	$7,060	$6,719	$341	90 (-34; 214)	1.3%
Combined (2019–2022)	$6,663	$6,367	$296	45 (-65; 155)	0.7%
Non-hospital spending and standardized hospital spending per beneficiary per year ($)[b]					
Baseline period (2017–2018)	$11,056	$10,626	$430		
MDPCP					
Combined (2019–2020)	$11,064	$10,725	$339	-91 (-272; 90)	-0.8%
Combined (2021)	$12,103	$11,762	$341	-89 (-305; 128)	-0.7%
Combined (2019–2022)	$11,422	$11,082	$340	-90 (-264; 84)	-0.8%
30-day post-discharge unplanned readmission (percentage)					
Baseline period (2017–2018)	15.9%	16.0%	-0.1%		
MDPCP					
Combined (2019–2020)	15.8%	16.0%	-0.2%	-0.1 (-0.8; 0.5)	-0.6%
Combined (2021–2022)	16.2%	16.5%	-0.3%	-0.2 (-0.9; 0.5)	-1.2%
Combined (2019–2022)	16.0%	16.3%	-0.3%	-0.2 (-0.8; 0.4)	-1.2%
Timely follow-up after acute exacerbation of a chronic condition (percentage)					
Baseline period (2017–2018)	75.2%	74.6%	0.6%		
MDPCP					
Combined (2019–2020)	74.3%	71.9%	2.4%	1.8** (0.6; 3)	2.5%
Combined (2021–2022)	74.3%	72.5%	1.8%	1.2 (-0.3; 2.6)	1.6%
Combined (2019–2022)	74.3%	72.2%	2.1%	1.5** (0.3; 2.7)	2.1%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[a] The percentage is calculated as the impact estimate for the year divided by the estimate of the counterfactual for the year. We estimated the counterfactual as the mean outcome observed that year for MDPCP minus the difference-in-differences impact estimate for the year.

[b] Spending is winsorized to the 99.5th percentile.

CI = confidence interval; ED = emergency department; MDPCP = Maryland Primary Care Program; PQI = Prevention Quality Indicator.

E.2. Methods for implementation findings

To learn about practices' and CTO experiences with implementing MDPCP, we analyzed primary data that we collected through virtual site visits to practices and CTOs and portal data submitted to CMS by MDPCP practices.

E.2.1. Primary data collection and analysis

We conducted one-time, virtual site visit interviews with 15 primary care sites and six CTOs from September 2022 to February 2023, which corresponds to the fourth and fifth year of MDPCP implementation for practices that started in 2019.

E.2.1a. Selection of practices and CTOs

We selected 15 practice sites, including 14 primary care practices that started MDPCP in 2019 and remained active in 2022 and one FQHC that joined in 2022 when FQHCs first became eligible for MDPCP[50]. We selected 15 practices that had characteristics that reflected the range of MDPCP practices that joined MDPCP in 2019 in terms of whether they partnered with a CTO, their 2022 Comprehensive Primary Care Payment (CPCP) percentage[51], level of HEART payments, and practice characteristics, including system affiliation, practice size, and region (Exhibit E.18). Three of the original practices selected in the sample declined to participate in interviews or were nonresponsive and were replaced with alternate practices with similar characteristics.

We selected six CTOs that had partnered with a practice in our sample. When identifying the six CTOs, we prioritized speaking to CTOs with diverse characteristics related to the number of practices and regions they served and the type of health care delivery organization, such as health system, physician-led, or other.

[50] Together with CMS, we decided to include one FQHC in the site visit sample to learn about the experience of this type of organization that was newly eligible to participate in MDPCP in 2021.

[51] CPCP percentage is defined as the percentage of the CPCP that practices opt to receive upfront. Remaining funds are distributed as applicable FFS payments for Evaluation & Management services.

Exhibit E.18. Characteristics of practices in site visit sample compared to all MDPCP practices that started in 2019 and remained active in 2022

Practice characteristic	Practices in sample[a] (n=14)	All practices (n=340)
CTO support[b]		
50/50	40%	49%
70/30	33%	30%
No CTO partner	27%	21%
CPCP percentage		
10%	20%	35%
25%	27%	27%
40%	27%	25%
65%	20%	13%
HEART payments		
1st quartile	13%	25%
2nd quartile	27%	25%
3rd quartile	33%	25%
4th quartile	20%	25%
System affiliation		
System-affiliated	47%	47%
Not system-affiliated	43%	53%
Size		
Small (1–2 providers)	27%	34%
Medium (3–5 providers)	40%	38%
Large (6 or more providers)	27%	28%
Region		
Baltimore City	20%	13%
Capital Region	20%	27%
Central Maryland	27%	35%
Eastern Shore	13%	9%
Southern Maryland	7%	6%
Western Maryland	13%	10%

Source: Mathematica's analysis of 2022 MDPCP participation data provided by The Lewin Group and OneKey[TM] data from IQVIA. The definitions for the regions come from the Maryland Hospital Association.

Notes: N = 340 primary care practices and FQHC organizations that were active in 2022 and joined MDPCP in 2019. Percentages for CPCP percentage, HEART payment, and size do not add up to 100%. Those data were not available for the FQHC in the sample.

[a] We conducted site visits to 14 practices and one FQHC. The FQHC's characteristics are not included in this table or anywhere else in the report to protect confidentiality.

[b] MDPCP practices can choose to partner with a CTO to support MDPCP implementation at their practices. Partnering practices can choose to defer either 50 or 30% of their CMF and HEART payments to the CTO. Practices that defer 50 % receive a higher level of support from CTOs, including care managers.

CMF = care management fee; CTO = Care Transformation Organization; CPCP = Comprehensive Primary Care Payment; FQHC = Federally Qualified Health Center; HEART = Health Equity Advancement Resource and Transformation; MDPCP = Maryland Primary Care Program.

E.2.1b. Conducting interviews

We conducted one-on-one telephone interviews with practice and CTO respondents. In a few cases, we conducted small group interviews with two or three respondents. For each practice in our sample, we interviewed up to six staff who work on MDPCP-related activities at the practice. Practice respondents included physician leaders, administrative leaders, and frontline staff (including physicians, nurse care managers, program coordinators, and quality improvement leads). For each CTO in our sample, we interviewed one leader who was familiar with the types of support MDPCP practices received from the CTO.

We developed two interview guides, one for practice respondents and one for CTO respondents. Both interview guides covered the following topics: reasons for participation in MDPCP, MDPCP's alignment with existing initiatives, the value of the different MDPCP payments and how they are used, the MDPCP data reports and learning supports, and effects of COVID-19 on care delivery changes. The practice interview guide also included questions on their perspectives on care delivery changes made since the first year of MDPCP participation or because of MDPCP, including implementation challenges and facilitators, transitioning through MDPCP tracks, working with CTOs, and influence of the Statewide Integrated Health Improvement Strategy goals on care delivery changes they made. The CTO interview guide included questions on the supports and services they provide to MDPCP practices.

E.2.1c. Analysis of interview data

We transcribed all interview recordings. The research team coded the interview data using the NVivo software, using codes that aligned with the topics covered in the interview guide. The research team analyzed the coded data by generating code reports by topic and practice then summarized the data segments in each code report in NVivo. For each topic, we reviewed the summaries across practices to identify cross-cutting themes.

E.2.2. Secondary data analysis

E.2.2a. MDPCP practice portal data

CMS requires practices to self-report their care delivery changes through the MDPCP practice portal. CMS uses these data to track practice progress on the care transformation requirements and to inform learning activities, as well as track practice compliance with the model. Practices completed the survey each quarter in 2019, in quarter 3 of 2020, and then twice a year (quarter 1 and quarter 3) in 2021 and 2022.

We examined care delivery changes over time for the 336 practices that started in 2019 and continued through the end of 2022. This approach enabled us to track the evolution of care delivery changes throughout the model period. For the 2019-starters, we used 2019 quarter 1 data as baseline data reflecting implementation at the outset of the model, and then we used the latest data available in each year to represent their experience in that year. We used 2019 quarter 3 or 4 data to represent 2019 (the latest quarter in which the given question was asked) and then 2020, 2021, and 2022 quarter 3 data to represent each subsequent year.

MDPCP practice portal data is useful for assessing care delivery changes across MDPCP practices because it includes the 336 practices that started in 2019 and continued to participate through 2022. A limitation, however, is that practices self-report the portal data and may have an incentive to over-report progress to CMS. We synthesized the MDPCP practice portal data with interview data to assess consistency across data sources and improve our understanding of practices' care delivery changes during MDPCP. For reference, Appendix E.2.4 includes a table summarizing practice portal data by year.

E.2.3. Key findings

E.2.3a Participation experience

Why practices and CTOs joined MDPCP

Nearly all practices we interviewed reported wanting to participate in MDPCP because they thought that the MDPCP requirements, as well as the payments, and/or CMS-sponsored learning supports would help provide better care for their patients. Many practices reported valuing the MDPCP care transformation requirements for establishing structure and standards that focus and guide practice change. Most practices also found that MDPCP payments enabled them to invest in resources necessary for practice change, such as paying for staff time or upgrading their electronic health record (EHR) systems. Finally, most practices said that the CMS-sponsored learning supports helped them implement care delivery changes, most commonly citing the practice coaches and virtual learning events as the most valuable learning supports. A couple of practices noted that practice coaches were a unique part of MDPCP who are not available in other value-based models and motivated these practices to join and remain in MDPCP.

Many practices also said that they joined MDPCP because of the perceived financial benefits in the short and long terms. In the short term, MDPCP provided practices immediate funding to invest in practice transformation activities. These practices believed that these short-term investments had long-term payoffs by strengthening the practice's capacity to deliver advanced primary care, which they hoped would prepare them to succeed in future value-based initiatives.

Nearly all system-owned practices we interviewed said that their systems made the decision to participate in MDPCP. This finding aligns with our findings that systems also drive MDPCP implementation at their practices, typically implementing system-wide changes for their MDPCP *and* non-MDPCP practices (such as Screening, Brief Intervention, and Referral to Treatment or advance care planning), as well as managing MDPCP reporting and payments for practices. The system-driven nature of MDPCP was particularly evident in a couple of system-owned practices where practice staff we interviewed were not familiar with MDPCP or with how their work related to the program. For independent practices, the physician owners made the decision to participate in MDPCP.

Overall experience with MDPCP

Most practices that participated in site visits valued their experiences in MDPCP and intended to continue in the program. The exceptions were two small independent practices, where the physician leaders reported finding the program's requirements to be overly burdensome for the level of payment they received.

Many practices reported concerns about transitioning to Track 3, specifically related to the lack of timely information on how care transformation requirements and payments will change in the new track. One system leader explained that health care organizations plan their budgets and programming years in advance, and they could not adequately plan or allocate resources for MDPCP without knowing how the level of payments and requirements will change. Many practices reported finding the payment methodology for Track 3 confusing, which prevented them from calculating changes to their revenue and financial viability in the model or as a business. A couple of these practices expressed frustration that CMS and MDPCP practice coaches had not answered their questions to clarify the changes to Track 3. Finally, a few practices said they purposely delayed transitioning to Track 3, because they wanted more time to gather information and prepare.

In contrast to the concerns about Track 3, many practices reported experiencing seamless transitions to Track 2. They reported feeling prepared and well supported for the transition by their MDPCP practice coaches, CTOs (for practices that partner with CTOs), and health systems (for practices affiliated with a health system). To prepare for Track 2, practice coaches monitored practices' progress in meeting the care transformation requirements needed to advance to Track 2 and helped them set and reach goals to meet the requirements, as needed. Furthermore, unlike the changes to the payment model in Track 3, practices did not perceive that the changes to the payment model for Track 2 would cause large disruptions to their revenue.

E.2.3b Payments

Practices' and CTOs' use and perceptions of the care management fee

Practices described the care management fee (CMF) as a unique and desirable component of the MDPCP payment model when compared to other value-based initiatives (like accountable care organizations). This is because the CMF is a large upfront payment that is not at risk. CMFs provided practices upfront financial capital to invest in practice transformation and improve their performance on quality and utilization measures a whole year before they are at risk for paying penalties.

Many practices thought that they received adequate funding to implement practice transformation changes required by MDPCP, because of the CMF. A few practices, however, were concerned that their level of payment in MDPCP was lower than it should be, due to what they described as problems with how beneficiaries were attributed to their practice. Only one solo independent practice said that MDPCP payments were not enough to make any improvements in care delivery at her practice. She noted the payments only covered the time she spent on MDPCP reporting instead of seeing patients.

Most practices reported using MDPCP funds, and the CMF specifically, to hire nurse care managers. Many practices also reported hiring behavioral health specialists and pharmacists, as well as social workers and community health workers who typically focused on connecting patients to resources that could help them meet their social needs. A few systems practices reported hiring practice transformation specialists who help practices with reporting, generating, and reviewing data reports, as well as conducting meetings with practice leaders to identify areas for performance improvement.

Health systems had the advantage of pooling MDPCP funding across practices to hire full-time staff who worked across multiple practices. This was almost always the case for pharmacists and behavioral health staff hired by health systems, who worked part-time across multiple practices, depending on the practice’s size and need. The ability to pool MDPCP funding to hire staff was a unique advantage for health systems. Independent practices, on the other hand, may only have enough funding for part-time staff who, as a couple of independent practices reported, are more difficult to recruit and retain.

A few practices said that MDPCP payments, and CMFs specifically, enabled care teams to spend more time on each patient—both during and between visits. This included letting doctors spend more time with certain patients during a visit and protecting staff time between visits for care management or calling patients to address gaps in care.

How practices and CTOs used and perceived the HEART payments

Most practices reported significant challenges using the HEART payments, even though they felt positively about the intent of HEART payments. Most practices appreciated how the HEART payments considered the extra support that practices need to give the most medically and socially vulnerable patients. However, many practices observed fundamental flaws in the criteria used to identify HEART-eligible patients and described their perceptions that the HEART eligibility criteria misidentified patients as vulnerable and missed patients who practices knew to be medically complex and socially vulnerable. Specifically, a few practice respondents said that zip codes cannot not accurately identify social vulnerability because certain zip codes have large economic diversity. One lead physician gave the example of urban areas with diverse housing: “There are places here with a lot of high rises and gated communities in neighborhoods that at one time were all low-income.” These practices expressed frustration that they could not reallocate HEART resources to patients they knew to need the extra support.

Many practices also expressed frustration with the restrictions on how HEART payments can be used. They described the administrative complexity of restricting who is eligible for services, especially when HEART eligibility changes quarterly. In addition, a couple of practices described the HEART payments as overly restrictive in how they can be used. This made it difficult for practices to provide HEART patients with resources they actually needed. On the other hand, as one practice reported, if they found a resource that the patient needed and could be covered by the HEART payment, the HEART payment did not always cover the full cost of the resource—which meant that either the practice or patient had to cover the remaining cost, unintentionally causing more financial burden on practices and patients. In general, practices believed that they provide more valuable and efficient care to their patients in programs that allow them to use their on-the-ground knowledge to inform who will receive what services.

Examples of HEART payment uses:

- Nutrition resources (medically tailored meals, grocery store gift cards)
- Transportation to medical appointments
- Equipment not covered by Medicare (like pulse oximeter, portable blood pressure cuffs, home-based modifications, etc.)
- Staff who support HEART-eligible patients (like community health workers) ◢

Many practices reported needing more time and clearer guidance from CMS to prepare for HEART payments. One system leader explained how it can take months, if not years, to design a new program or service and determine an effective way to implement it throughout their organization. Moreover, practices found the initial guidance about HEART payments from CMS unclear and vague. The leader of one CTO said, “It’s been hard to feel like you’re using the payment the right way because they’re not giving a lot of specifications.” As a result, these practices doubted how effectively they used HEART payments to support their patients.

On the other hand, a few practices and health systems reported making valuable changes using the HEART payments. These practices and health systems typically had larger levels of HEART payments, which they typically used to hire staff to support HEART-eligible patients. For example, a large health system we interviewed reported hiring community health workers who serve HEART patients across all its MDPCP practices, helping connect patients to resources that can help meet their social needs.

How practices and CTOs used and perceived the performance-based incentive payment and CPCP

Practices did not make big changes using the performance-based incentive payment (PBIP) and CPCP, because the PBIP is at risk and both payments are a small portion of the total MDPCP payments given to practices. A couple of practices said that they do not consider the PBIP a part of their revenue, because they might need to pay it back. The CPCP is not at risk but, like the PBIP, is a small amount compared to the CMF and HEART payments and so does not influence practice investments, according to a few practices.

Many practices already had processes in place to improve their performance on measures that impact the PBIP. These practices said that the PBIP measures align with measures from other value-based initiatives, so the measures did not drastically change their focus or care delivery. Most practices reported relying on existing staff and infrastructure to improve their performance on quality measures, such as quality improvement specialists, as well as monthly score cards and quality workgroups. A few health systems and large independent practices reported making new investments with MDPCP funds to improve their PBIP performance—such as quality improvement or practice transformation specialists, Healthcare Effectiveness Data and Information Set nurses, and in-house services like labs and retinal eye exams to improve access to care for patients with diabetes.

Practices were mixed on whether the PBIP affected provider salaries. A few practices said that a provider's performance on measures related to the PBIP was factored into their compensation, while other practices said that the PBIP was so small that they did not factor it into provider compensation.

Many practices reported concerns about the fairness or appropriateness of the measures used to calculate PBIP amounts. They said that the cost and utilization measures are not only affected by primary care, and that some of the quality measures (such as patient experience of care and body mass index screening and follow-up plan) do not correlate with better care and are a waste of their time.

E.2.3c Care delivery changes

The MDPCP model required participating practices to make care delivery changes across five functions of primary care: (1) access and continuity, (2) care management, (3) comprehensiveness and coordination, (4) patient and family engagement, and (5) planned care and population health. Below, we list the care delivery requirements for MDPCP practices and describe the changes that MDPCP practices made to meet these requirements based on interview and portal data from practices that started participating in MDPCP in 2019 and remained enrolled through 2022. Appendix E.2.4 includes a supplemental table showing practices' responses to MDPCP practice portal items by year.

Function 1: Access and continuity

Empanelment. CMS required MDPCP practices to empanel patients to a practitioner or care team as a first step towards improving the continuity of care.

– Practices' efforts to empanel patients pre-dated MDPCP. Since the outset of MDPCP, the median percentage of beneficiaries who were empaneled to a practitioner or care team at MDPCP practices was 100%, according to practices' reports to CMS.

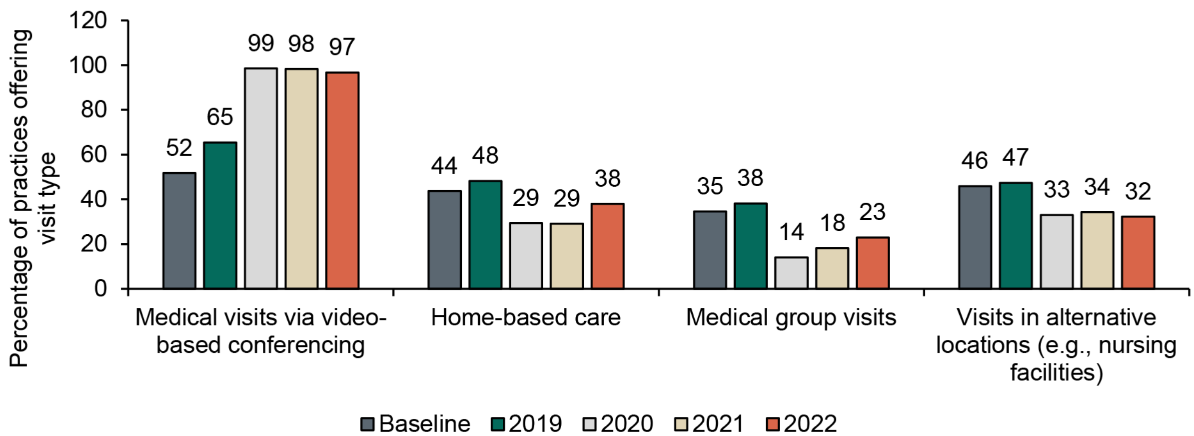
24/7 access to a care team or practitioner. CMS required MDPCP practices to ensure that MDPCP beneficiaries had 24/7 access to a care team or practitioner with real-time access to the EHR.

- Most practices implemented 24/7 access before MDPCP began. Since the outset of the model, nearly all MDPCP practices reported to CMS that they offered 24/7 access for all patients at the practice (increasing from 95% of practices at the beginning of 2019 to more than 99% of practices from 2020 through 2022).

Alternative care (Track 2 requirement). Track 2 MDPCP practices were required to use MDPCP alternative payments to ensure that beneficiaries had access to their practitioner or care team through at least one alternative care strategy outside of FFS office visits (such as telehealth, group visits, or home visits).

- Practices increasingly offered visits via video-based conferencing but reported limited progress implementing other types of alternative visits, such as home visits or group visits (Exhibit E.19). During site visits, a few practices said they had planned to offer group visits and home visits for MDPCP but paused these efforts when the COVID-19 pandemic began.
- Practitioners and staff at practices that participated in site visits attributed adoption of telehealth to COVID-19 and FFS payments made available for telehealth during the public health emergency rather than to MDPCP.

Exhibit E.19. The percentage of MDPCP practices that offered medical visits via video-based conferencing increased in 2020, while provision of other alternative visit types remained relatively low



Notes: Mathematica's analysis of MDPCP practice portal data from baseline (N = 336), 2019 (N = 336), 2020 (N = 336), 2021 (N = 336) and 2022 (N = 332).

Other changes to improve access and continuity. Although not required, CMS encouraged practices to make other changes to improve MDPCP beneficiaries' access to and continuity of care. For example:

– *CMS encouraged MDPCP practices to expand hours of operation; while more practices reported offering after-hours and weekend appointment availability, some practices may not have implemented this as the model had intended.* The percentage of practices that reported to CMS that they often or always offered office visits after hours or on weekends increased by 23 percentage points over the course of MDPCP (from 52% at the outset of the model to 70% at the end of 2019, increasing to a high of 75% at the end of 2022). However, practices may have used inconsistent definitions of "expanded hours". A few system-owned practices that participated in site visits noted that they had implemented expanded hours by referring patients to off-site system-owned urgent care facilities. In addition, a few practices (mainly smaller practices) said they did not implement expanded hours due to lack of demand from patients and/or staffing issues.

– *CMS encouraged practices to offer same-day or next-day appointments, but almost all practices reported doing this before MDPCP began.* At the outset of the model, 93% of practices reported that they often or always provided same- or next-day appointments to beneficiaries who needed them. This percentage increased to 99% by 2021.

Function 2: Care management

Risk stratification. To identify patients with the greatest needs, MDPCP practices were required to ensure that all empaneled beneficiaries were risk stratified.

– While most practices risk stratified patients prior to MDPCP, practices reported to CMS that they enhanced their risk stratification approaches during MDPCP. For example, the percentage of practices that reported using a two-step risk stratification method[52] increased gradually from 47% at the outset of MDPCP to 75% in 2022 (a 28-point increase). The percentage of practices that reported integrating risk stratification into their EHR or health information technology (IT) system also increased gradually from 49% at the outset of MDPCP to 64% in 2022 (a 15-percentage point increase).

– System- and practice-level staff had differing perspectives on risk stratification. During interviews, system and CTO leads from many practices described improving the accuracy of risk scores during MDPCP by using the Hilltop Pre-AH Model reports[53] from Chesapeake Regional Information System for our Patients (CRISP) and other data. In

[52] Two-step risk stratification uses algorithms based on patient risk factors and clinical judgement to assign and adjust patients' risk scores.

[53] The Hilltop Pre-AH Model is a risk prediction model that uses a variety of risk factors derived from Medicare claims data to estimate the probability that a given patient incurs an avoidable hospital event in the near future. For additional information, see https://health.maryland.gov/mdpcp/Documents/The%20Hilltop%20Pre-AH%20Model%20In%20Brief.pdf#:~:text=The%20Hilltop%20Pre-AH%20Model%E2%84%A2%20is%20a%20risk%20prediction,incurring%20an%20avoidable%20hospitalization%20or%20emergency%20department%20event.

contrast, practitioners and staff from many practices said that risk scores were not useful, either because they preferred to rely on clinical judgment or they perceived risk scores to be inaccurate or outdated.

Care management services. CMS required practices to provide (1) targeted, proactive, relationship-based (longitudinal) care management services to beneficiaries who are at an increased risk and are likely to benefit from these services and (2) short-term (episodic) care management services to beneficiaries with a triggering event, such as a new serious illness or injury, major life event, newly unstable chronic illness, or transition to a new care setting.

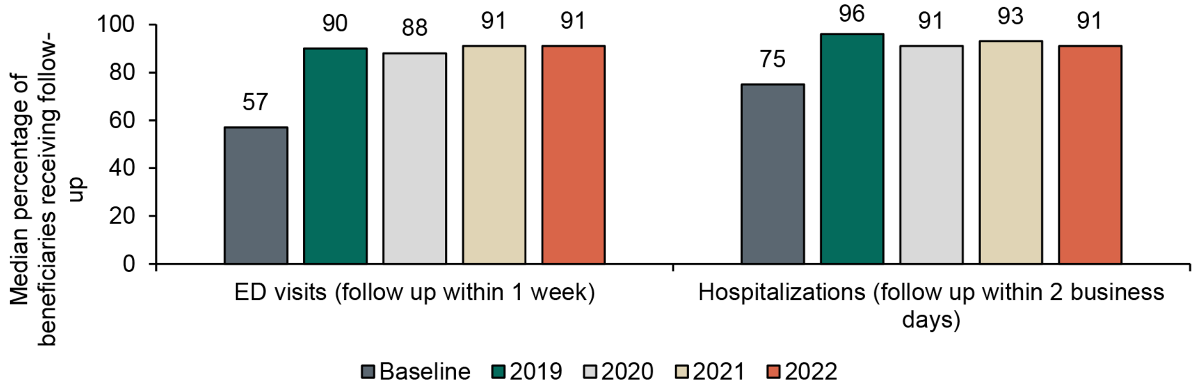
- Practices increasingly hired care managers during MDPCP. The percentage of practices that reported having a designated care manager (employed by the practice or CTO, including on-site and off-site care managers) increased by 9 percentage points from 85% in 2020 (the first year the question was asked) to 94% in 2022.
- The median percentage of all empaneled beneficiaries under care management (including longitudinal and episodic) increased gradually from 1% at the outset of the model to 14% in 2022.
- During site visits, all practices said they hired care management staff with MDPCP funds or partnered with CTO care management teams. Most practices said care management teams typically consisted of registered nurses (RNs) and/or social workers but occasionally included other types of staff, including community health workers, diabetes educators, or behavioral health specialists. Similar to RN care managers, care managers without a nursing background conducted proactive outreach to high-risk patients, but their main responsibility was to listen to patients' needs and link patients to services and supports, such as visits with primary care practitioners and diabetes educators or supports for health-related social needs (HRSNs).
- Most practices that participated in site visits said they formalized care management workflows and/or dedicated additional staff to provide longitudinal care management because of MDPCP, making it more systematic and widespread.
- Care managers supported a range of needs, including short-term support for new diagnoses, support meeting longer-term goals (like reductions in blood pressure or A1c), and referrals to other medical and service providers (including specialists, behavioral health providers, or community-based organizations). At a few practices, care managers gave patients their phone numbers if they had questions or concerns between visits.
- Most practices said the changes they made to care management offerings for MDPCP improved a variety of outcomes. For example, they said it improved patients' access to care by providing a direct point of contact for questions, prevented exacerbation of chronic conditions and reduced hospitalizations through regular contact with patients, and improved patient satisfaction by helping patients feel "seen". "[Care managers] have been able to intervene...and handle something from an outpatient perspective that would have otherwise required an inpatient [stay]," said one care manager.

- Many practices said that having care managers embedded in the practice was beneficial for improving coordination and trust among patients, practitioners, and care managers (versus working with off-site, remote care managers).
- However, a few practices noted that care management for patients with long-term and short-term needs was inadvertently de-prioritized due to care managers' competing responsibilities to conduct timely follow-up with patients discharged from the hospital or emergency department (see below). As one care manager noted, "The priority of [care management] gets downgraded because of the other more higher-priority things that get put on the top of the list." A few practices reported that they had to limit the number of patients they could provide ongoing care management services to and thus were unable to reach all patients who would benefit. For example, one practice reported that they could only serve 5% of high-risk patients and missed opportunities to work with many rising-risk patients.

Follow up after emergency department or hospital discharges. MDPCP practices were required to ensure that all MDPCP beneficiaries received a follow-up interaction from the practice within one week for emergency department discharges and two business days for hospital discharges.

- Practices increasingly tracked discharges from the hospital and emergency department. The percentage that reported tracking discharges from the emergency department and hospital doubled during the first two years of the model (from around 50% at the beginning of 2019 to 100% at the end of 2020, where it remained through 2022).
- Discharge follow-up rates improved during MDPCP (Exhibit E.20). Among the roughly 50% of practices that tracked emergency department discharges at baseline, the median percentage of beneficiaries receiving follow-up within one week increased by 35 percentage points (from 57% at the beginning of the model to 90% at the end of 2019, where it stabilized through 2022). Among practices that tracked hospital discharges at baseline, the median percentage of beneficiaries receiving follow-up within two days increased by about 20 percentage points (from 75% at the outset of the model to 96% at the end of 2019, before stabilizing at 91% through 2022).

Exhibit E.20. Among MDPCP practices that tracked discharges at baseline, practices reported an increase in the median percentage of beneficiaries receiving a timely follow-up call after ED and hospital discharge



Notes: (1) For hospitalizations: Based on Mathematica's analysis of the MDPCP practice portal data from baseline (N = 174), 2019 (N = 169), 2020 (N = 171), 2021 (N = 173), and 2022 (N = 159) among the 180 practices that reported tracking hospitalizations at baseline. (2) For ED visits: Based on Mathematica's analysis of the MDPCP practice portal data for baseline (N = 156), 2019 (N = 159), 2020 (N = 158), 2021 (N = 157) and 2022 (N = 171) among 174 practices that reported tracking ED visits at baseline

ED = emergency department; MDPCP = Maryland Primary Care Program.

- Most practices that participated in site visits described making major changes to staffing and workflows to track discharge data and ensure patients received timely follow-up after hospital and emergency department discharge. Most practices had designated care managers call patients to review discharge instructions, discuss next steps (such as scheduling primary care appointments), and ensuring patients had medications. Practices used MDPCP funds to hire care managers who conducted discharge follow-up and many said they used CRISP to identify patients needing follow-up.
- At a few system-owned practices that participated in site visits, a team of care managers at associated hospitals (rather than practice-level care managers) conducted discharge follow-up calls with patients. A couple of other system-owned practices mentioned that care managers from both the practice and system conducted discharge follow-up calls and that these efforts were not well coordinated; sometimes, patients received discharge follow-up calls from both care managers.
- Many practices said that discharge follow-up efforts helped reduce readmissions and improve patient outcomes post-discharge. A few respondents said these changes were the most valuable changes they adopted for MDPCP. For example, one respondent said, "With [discharge follow-up], we have been able to reduce our readmission rates, we've been able to keep people out of the hospital. And if they do have, for whatever reason, chronic readmissions...we've been able to reduce the stay in the hospital, which has helped improve outcomes in terms of complications, fatalities."

- However, many practices that participated in site visits also cited challenges to discharge follow-up, especially lack of time, difficulty reaching some patients, and challenges using CRISP (primarily due to data bugs and lags).

Care planning (Track 2). CMS required Track 2 practices to ensure that MDPCP beneficiaries in longitudinal care management were engaged in a personalized care planning process, focused on their goals, needs, and self-management activities.

- While nearly all practices (including those in Track 1 and Track 2) reported documenting and storing care plans before MDPCP began, practices reported expanding the number of beneficiaries for whom they used care plans during MDPCP. Each year, approximately 90% reported that care plans are integrated in the EHR or other health IT; fewer than 2% of practices said they did not document and store care plans at the outset of the model. However, from 2019 to 2022, the percentage of practices (including those in Track 1 and 2) that reported using care plans for most or all beneficiaries under longitudinal care management increased from 53% at the outset of the model to approximately three-quarters of practices in 2022.
- A few practices that participated in site visits described using care plans for at least some patients in longitudinal care management. Of these, most said they used disease-specific templates for common chronic conditions (diabetes, hypertension, etc.), embedded in their EHR. A couple of practices cited challenges to using care plans, including difficulty accessing and searching for information entered in care plan templates in the EHR and difficulty engraining use of care plans in workflows.

Comprehensive medication management (Track 2). CMS required Track 2 practices to ensure that beneficiaries in longitudinal care management had access to comprehensive medication management (CMM).

- The percentage of practices that reported to CMS that they provided or facilitated access to CMM services for beneficiaries increased by 44 percentage points (from 55% to 99%) over the course of the model. Among practices that participated in site visits, most practices in CTOs reported that their CTO established formal agreements with pharmacists to provide CMM services to MDPCP beneficiaries. Practices typically said practitioners were responsible for referring patients to the pharmacist for CMM services, although one practice said that the care manager reviewed risk score reports to proactively identify patients for CMM services.
- Despite increased access to CMM services at MDPCP practices, many practitioners whose practices partnered with off-site pharmacists for CMM noted that they rarely referred patients to these pharmacists, either because they forgot about the availability of CMM services or because they believed it was more effective to manage patients' medication on their own.

Function 3: Comprehensiveness and coordination

Referral management. CMS required MDPCP practices to ensure coordinated referral management for MDPCP beneficiaries seeking care from high-volume and/or high-cost specialists, as well as emergency departments and hospitals.

- Many practices enhanced coordination of referral management with specialists in 2019 and continued to do so throughout MDPCP. The percentage of practices that reported coordinating referral management with specialists increased from 51% at the outset of MDPCP to 93% at the end of 2019. From 2020 through 2022, the percentage of practices that coordinated referral management continued to increase to more than 99%. Many practices that participated in site visits described tasking care managers or medical assistants with helping patients schedule appointments with specialists, following up with specialists for documentation, or addressing patients' barriers to attending specialist referrals (such as transportation). However, only a few practices reported making these changes during MDPCP—most said their efforts predated 2019.
- The model encouraged practices to review data on high-volume and high-cost specialists, but practices generally did not focus their efforts on this. Only two practices that participated in site visits mentioned reviewing data on high-cost or high-volume specialists, and both practices noted that patients tend to go to specialists that they want to see, regardless of PCPs' referral suggestions.

Behavioral health integration. MDPCP practices were required to ensure that beneficiaries with behavioral health needs had access to behavioral health care using an approach for integration suggested by the model. Suggested approaches include evidence-based behavioral health integration models, such as the Primary Care Behaviorist (PCB) model, which requires practices to co-locate behavioral health professionals in the primary care setting to address patients' behavioral health needs, and the Collaborative Care Model (CoCM), which requires practices to offer relationship-based care management from a care manager with behavioral health training, as well as coordinated referrals to a behavioral health provider (for patients whose needs exceed the scope of primary care).

- The percentage of practices that used one of these behavioral health integration models increased slightly during MDPCP, while approximately one-third of practices continued to rely on external referrals for behavioral health needs. From the outset of MDPCP through 2022, the percentage of practices that reported using either the PCB model or CoCM to address patients' behavioral health needs in the primary care setting increased from 53 to 62%. An additional 29 to 38% of practices reported annually they did not use the PCB model or CoCM, but instead referred patients with behavioral needs to external providers.
- The proportion of patients whose behavioral health needs were addressed in the primary care setting increased slightly during MDPCP. Among practices that reported using the PCB model or CoCM, the percentage that reported addressing behavioral health needs for

most or all patients with behavioral concerns increased slightly from 25% at the outset of the model to 32% in 2022, while the percentage that reported addressing needs for at least some patients increased from 66% to 100% during the same period.

– Among practices that participated in site visits, those that integrated behavioral health services in the practice setting described this change as valuable for linking beneficiaries to prompt care for behavioral health needs. Although many practices said they continued to rely on external referrals to behavioral health providers, a few described embedding a behavioral health specialist in the practice one or more days per week, noting that they decided to do so because of the MDPCP requirements and availability of MDPCP funds. These practices unanimously noted that this was a valuable change that helped patients get prompt support for behavioral health needs in spite of widespread provider shortages and months-long waitlists to see behavioral health providers.
– A few practices that participated in site visits said their CTO provided shared off-site behavioral health specialists or access to Mindoula, a telehealth counseling platform. However, these practices noted that the resources were less useful because they were not truly embedded in the practice; therefore, it was harder for practitioners to coordinate with these services, and patients were less trusting of them.
– In addition, practices increasingly reported implementing the Screening, Brief Intervention, and Referral to Treatment (SBIRT) in 2022. From 2021 (the first year the question was asked) to 2022, the percentage of practices that reported using SBIRT increased from 51% to 64%. Many practices that participated in site visits described system- or CTO-wide efforts to educate practices on the use of SBIRT, but a few practice respondents highlighted noteworthy challenges, including time constraints and limited treatment options for patients identified as needing additional support.

HRSNs (Track 2). Track 2 practices were required to facilitate access to resources in the community for MDPCP beneficiaries with identified HRSNs.

– Practices increasingly screened patients for HRSNs and maintained inventories of social service resources. The percentage of practices (including Track 1 and Track 2) that reported screening all patients for unmet social needs increased gradually from 21% at the outset of MDPCP to 64% in 2022 (a 43-percentage point increase), and the percentage that reported maintaining an inventory of social service resources increased from 5% at the outset of the model to 98% or greater in performance years 2 through 4 (a 94-percentage point increase).
– Most practices that participated in site visits said they had workflows in place to assess and address patients' HRSNs, but about half of these practices said these efforts predated MDPCP. Generally, these efforts included assigning responsibilities to address HRSNs to care managers or community health workers or to making regular updates to resource inventories.
– A few practices used HEART payments and CTO resources to address HRSNs. They described using HEART payments to fund the provision or expansion of services (such as

resources to address food insecurity). A few practices that partnered with CTOs said their CTOs provided access to care managers and social workers who were available to work with patients to address HRSNs.

– However, practices faced challenges linking patients to resources to address HRSNs. Many practices reported limited availability of community resources to refer patients, especially transportation resources. A few other practices said that staff assigned to address HRSNs did not have enough time to focus on this part of their job.

Function 4: Patient and family engagement

Patient and Family Advisor Councils (PFACs). CMS required practices to convene PFACs at least annually and integrate PFAC recommendations into care and quality improvement activities.

– Most practices implemented PFACs for the first time in 2019 and continued to hold PFACs at least annually. At the outset of the model, 18% of practices reported holding a PFAC during the past 12 months. This increased to 74% by the end of 2019 and continued to increase to over 90% in 2021 and 2022 (the question was not asked in 2020).

– Practices that participated in site visits shared mixed feedback on the effectiveness of PFACs. A few reported that PFACs helped them make changes that improved patient experience (such as reducing wait times and improving the office telephone system that patients use when calling to schedule appointments). A few others described challenges to scheduling PFACs, recruiting new patients with diverse perspectives, and collecting actionable feedback.

Advance care planning (Track 2). Track 2 practices were required to engage MDPCP beneficiaries and caregivers in a collaborative process for advance care planning (ACP).

– Almost all practices (98%) reported providing ACP before MDPCP, but practices enhanced delivery of ACP during MDPCP. For example, practices reported greater communication with beneficiaries and proxies about care preferences (increasing from 71% when the question was first asked in 2020 to 87% in 2022) and greater involvement of non-practitioner clinical staff in ACP (from 50% at the outset of the model in 2019 to 75% in 2022). Practices also increasingly reported storing ACP documents in patient portals or health records, although there is still room for improvement (from 26% at baseline to 39% in 2022).

– During site visits, a few practices (all system-owned) echoed this finding, noting that they had enhanced their approach to ACP during MDPCP by training practice staff and implementing workflows to systematically conduct ACP with patients. These practices noted that before MDPCP, individual practitioners were responsible for ACP, but there were no system-wide, coordinated efforts to make sure this was done consistently.

Function 5: Planned care and population health

Continuous improvement. CMS required practices to continuously improve their performance on key outcomes, including cost of care, electronic clinical quality measures, beneficiary experience, and utilization measures.

– Although nearly all practices reported focusing on at least one measure to guide quality improvement since the outset of the model, practices expanded their use of data and measures for quality improvement during MDPCP. For example, the percentage of practices that reported meeting and reviewing quality improvement data at least monthly increased gradually from 45% at the outset of MDPCP to 79% in 2022 (a 34- percentage point increase). Practices also reported increases in the availability of electronic clinical quality measure data, claims data feedback from CMS, beneficiary experience data, patient-reported outcome measures, multi-payer data, public health data from the county or state, and internal practice or system data.

– To discuss planned care and population health, practices increasingly relied on care team meetings. The percentage of practices that reported holding scheduled care team meetings at least weekly to discuss high-risk beneficiaries and planned care increased by 30 percentage points, from 15% at the outset of the model to 45% in 2022.

– Most practices that participated in site visits described increasing their focus on planned care and population health during MDPCP. For example, many practices said they shifted or increased focus on MDPCP measures used to calculate the PBIP (such as emergency department and hospital utilization), implemented new workflows to conduct outreach to close gaps in care, and increased the frequency that they reviewed data and reports. A few practices said they hired new staff, such as medical assistants and nurses, to support this work, although it is unclear whether MDPCP funds were used to support these salaries. Practices that partnered with CTOs described some CTO involvement, ranging from distributing dashboards summarizing key measures to providing technical assistance to improve the accuracy of measures (for example, by helping practices enter data in the EHR more consistently or updating templates in the EHR to improve data entry).

– During site visits, practices shared mixed feedback on the value of using data for continuous improvement. Many practices said using data on quality and utilization helped patients avoid “falling through the cracks.” A few others noted that the enhancements to team-based care related to MDPCP (such as hiring care managers or other staff using MDPCP funds) helped ensure patients’ needs were met. However, practices also described challenges using measures to improve care. For example, many practices noted that enhanced focus on data and measures increased administrative burden and decreased the amount of time practitioners had to spend with patients. A practitioner at one practice that participated in site visits noted, “The [emphasis on measures] is exhausting, its inefficient, its imprecise. Half the time I’m running to get and document information instead of focusing on the patient.” Many other practices said it is hard to improve on some quality and utilization measures because some patients do not comply with the practice's recommendations to get routine preventative care.

E.2.4. Supplemental Results

Exhibit 21. Table of practice portal data reported by MDPCP practices by year

= 1 or more percentage point decrease
= 11–39 percentage point increase
= 0–10 percentage point increase
= 40 or greater percentage point increase

MDPCP Practice Portal Item	Baseline	2019	2020	2021	2022	Change[a]
Access and continuity						
Median percentage of beneficiaries empaneled	100	100	100	100	100	**0**
Percentage of practices providing 24/7 coverage with real-time access to the electronic health record	95	93	99	100	100	**5**
Percentage of practices providing *home-based care*	44	48	29	29	38	**-6**
Percentage of practices providing *medical group visits*	35	38	14	18	23	**-12**
Percentage of practices providing *medical visit via video-based conferencing*	52	65	99	98	97	**45**
Percentage of practices providing *visits in alternative locations*	46	47	33	34	32	**-14**
Percentage of practices *always or often* able to provide office visits on the weekend, evening, or early morning when beneficiaries need it	52	70	80	74	75	**23**
Percentage of practices *always or often* able to provide same or next-day appointments when beneficiaries need it	93	95	96	99	99	**6**
Care management						
Percentage of practices that risk stratify empaneled beneficiaries	91	99	100	100	100	**9**
Percentage of practices using two-step risk stratification among practices that risk stratify empaneled beneficiaries	47	70	75	87	75	**28**
Percentage of practices with risk stratification integrated into the electronic health record or health information technology system	49	44	56	64	64	**15**

MDPCP Practice Portal Item	Baseline	2019	2020	2021	2022	Change[a]
Percentage of practices having a designated care manager employed by the practice or CTO	NA	NA	85	95	94	**9**
Median percentage of empaneled beneficiaries under care management at MDPCP practices	1	5	11	12	14	**13**
Percentage of practices that track discharges from the emergency department	52	89	100	100	100	**48**
Median percentage of beneficiaries receiving follow-up within one week, among practices that tracked emergency department visits at baseline (N = 174)	57	90	88	91	91	**34**
Percentage of practices that track discharges from the hospital	54	90	100	100	100	**46**
Median percentage of beneficiaries receiving follow-up within two days, among practices that tracked discharges from the hospital at baseline (N = 180)	75	96	91	93	91	**16**
Percentage of practices integrating care plans into the electronic health record or other health information technology	88	90	90	92	91	**3**
Percentage of practices using care plans for *most or all* beneficiaries under longitudinal care management	53	68	63	75	73	**20**
Percentage of practices providing or facilitating access to comprehensive medication management services for beneficiaries	55	65	82	94	99	**44**
Comprehensiveness and coordination						
Percentage of practices that have coordinated referral management with at least one high-frequency and/or high-cost specialty care provider	51	93	97	100	100	**49**
Percentage of practices using the Primary Care Behaviorist model or the Collaborative Care Model for behavioral health integration	53	58	60	60	62	**9**

MDPCP Practice Portal Item	Baseline	2019	2020	2021	2022	Change[a]
Percentage of practices addressing behavioral health needs for *most or all* patients with behavioral concerns, among practices using the Primary Care Behaviorist model or the Collaborative Care Model	25	27	16	25	32	**7**
Percentage of practices addressing behavioral health needs for *at least some* patients with behavioral concerns, among practices using the Primary Care Behaviorist model or the Collaborative Care Model	66	89	96	100	100	**34**
Percentage of practices that use Screening, Brief Intervention, and Referral to Treatment (SBIRT) to address behavioral health needs	NA	NA	NA	51	63	**12**
Percentage of practices routinely screening *all* patients for unmet social needs	21	36	48	57	64	**43**
Percentage of practices that have an inventory of social service resources	5	19	98	99	100	**95**
Patient and family engagement						
Percentage of practices that held a Patient and Family Advisory Council during the last 12 months	18	74	NA	93	97	**79**
Percentage of practices providing advance care planning	98	96	100	100	100	**2**
Percentage of practices that promote communication between beneficiaries and health care proxies regarding the beneficiary's values/ goals/ care preferences at the end of life as part of advance care planning	NA	NA	71	85	87	**16**
Percentage of practices that typically involve other clinical staff (RN, LPN, MA, care manager) in advance care planning	50	57	62	75	75	**25**
Percentage of practices that document and store advance care planning conversations and decisions in the patient portal/ patient health record	26	27	45	46	39	**13**

MDPCP Practice Portal Item	Baseline	2019	2020	2021	2022	Change[a]
Planned care and population health						
Percentage of practices meeting and reviewing quality improvement data (e.g., data on quality measures, cost, utilization, beneficiary experience of care) *at least monthly*	45	54	63	70	79	**34**
Percentage of practices having scheduled care team meetings *at least weekly* to discuss high-risk beneficiaries and planned care	15	28	38	40	45	**30**

Source: Mathematica's analysis of MDPCP practice portal data reported by MDPCP practices.

Notes: (1) N = 336 primary care practices that began participating in MDPCP at the start of 2019 and were active in 2022. The sample size for some items is slightly lower due to item-level non-response. (2) We defined the time periods as follows: Baseline = the first quarter in 2019 that the question was asked (some questions were asked in quarter 1, while others were asked in quarter 2); 2019 = quarter 3 or 4 of 2019 (the latest quarter in which a given question was asked in that year); 2020 through 2022 = quarter 3 of each year.

[a] Change is defined as the percentage point difference between 2022 and baseline (or the first year an item was asked if no baseline data is available).

NA = not available (question was not asked).

Appendix F. Methods and Supplemental Results for Estimating the Likely Spending Effects of Switching Maryland to the Prospective Payment System

F.1. Research design

Our research design assumed that, if Maryland switched to the Medicare inpatient and outpatient prospective payment system (PPS), per-capita spending and service use patterns in the state would shift—in the long term—toward patterns in similar geographic areas in the rest of the country that have long been operating under PPS. Our design, therefore, centered on developing national benchmarks, made up of matched groups of Public Use Microdata Areas (PUMAs) drawn from across the United States but outside Maryland. This strategy allowed us to compare actual Medicare per-capita spending and service use outcomes in Maryland in 2022 with a range of values for the same outcomes based on national benchmarks we constructed to represent Maryland operating under PPS.

Critically, our approach for developing benchmarks was meant to account for behavioral responses by hospitals and other providers that are likely to influence spending and service use changes in Maryland following a shift to PPS. Switching to PPS would lower Medicare prices substantially for hospital care, but looking at price effects alone (for instance, using claims in Maryland priced according to PPS rules) would not accurately identify how Medicare spending might change. This is because hospitals and other providers would likely respond to lower payment rates in ways that offset some apparent savings to Medicare—for example, by increasing volumes of care to regain some of the lost revenue and margins after switching from global budgets to PPS. Therefore, we prioritized constructing national benchmarks by matching PUMAs that are similar to Maryland on some characteristics, such as health status and demographics, and intentionally not matching on characteristics likely to change as a result of switching to PPS, such as the outcomes we examine.

F.2. Identifying likely behavioral responses from shifting to PPS

Before matching, we focused on identifying characteristics that are likely to change in response to Maryland shifting to PPS. In addition to informing decisions on whether to match on certain characteristics when constructing the national benchmarks, identifying likely behavioral responses also helps to explain *why* Maryland would move toward benchmarks after shifting to PPS.

We identified likely behavioral responses, shown in Exhibit 6.1, and described them in section 6.1 based on our understanding of current incentives of the Maryland Total Cost of Care (MD TCOC) Model, a review of the related literature, and interviews with six experts of the Maryland health system. There is a rich literature that shows hospitals exhibit behavioral responses to changes in payment policies. For example, Dafny (2005) finds hospitals often respond to Medicare price changes by "upcoding" patients to diagnosis codes that have higher rates. However, the literature that focuses *specifically* on behavioral responses from switching from an all-payer rate setting or global budget system to PPS is limited. Most related to our research question, the Centers for Medicare & Medicaid Services (CMS) noted potential increases in coding intensity, changes in hospital designations, and shifts of care from the hospital to post-acute settings were Maryland to shift to PPS in a report written before the start of the Maryland Total Cost of Care (MD TCOC) model (CMS 2018).

We supplemented our literature review with qualitative data gathered from a series of virtual interviews with six experts of the Maryland health care system completed in January 2022. To maintain confidentiality, we do not disclose the names of the experts. The experts did agree to disclose the organizations they are affiliated with, including the Commonwealth Fund, the Heath Services Cost Review Commission, Johns Hopkins, Mathematica, the Maryland Hospital Association, and the Urban Institute. As part of the interview protocol, the experts also agreed that we could use the qualitative information they shared to inform our quantitative research design and report on their responses anonymously. We also emphasized that the interviews are being conducted as part of a comprehensive independent evaluation, not because CMS is planning to switch Maryland to PPS.

The interviews were one hour and included the following questions:

1. What are the most important changes hospitals and nonhospital providers would make that would impact Medicare spending if Maryland converted from its current financing system to PPS?
2. What would be the timing of when these changes occur and the magnitude of their effects on Medicare spending?
3. What are likely implications for access, equity, and quality if Maryland switched to PPS?

F.3. Developing national benchmarks

We developed the national benchmarks in five steps, each of which is described in more detail below:

1. Selected PUMA as unit of analysis for matching
2. Identified and prioritized matching variables and set criteria for sufficient balance
3. Matched each PUMA in Maryland to PUMAs across the country to create the overall national benchmark group
4. Split the overall national benchmark group into high- and low-spending groups based on the median
5. Assessed the quality of the benchmark groups in terms of balance, size, and geographic distribution

F.3.1. Selecting PUMAs as unit of analysis for matching

Although our estimates in Chapter 6 reflect how Medicare spending and service use would change in Maryland *as a whole* after shifting to PPS, we selected PUMAs as the unit of matching. PUMAs, which are defined by the U.S. Census Bureau, are large enough units to limit variation in characteristics attributable to random noise but small enough to capture meaningful variation across geographic areas in Maryland. Each PUMA contains about 100,000 people and is constructed from census tracts and counties. In the 2010 Census, there were 44 PUMAs in Maryland, and 2,307 PUMAs in the rest of the country. Larger counties such as Baltimore City (a county equivalent) are divided into multiple PUMAs, enabling finer resolutions for key characteristics that vary within the county, while sparsely populated counties are combined into a single PUMA to help ensure reliability.

F.3.2. Identifying matching variables and setting criteria for what counts as sufficient balance

In close collaboration with CMS, we set priorities (high versus low) for matching variables based on associations with outcomes (conceptual and empirical) and face validity to improve matching feasibility (summarized in Exhibit F.1). In addition to setting priorities for matching characteristics, we set criteria for sufficient balance for each of the matching variables. Among the high-priority variables, we set hierarchical condition category (HCC) score, Regional Price Parity, and the percentage of Black residents as highest priority because these variables were found to be most strongly associated with Medicare spending levels.[54]

For all high-priority variables, we targeted standardized differences between Maryland and the benchmark group of less than 0.15. For the low-priority matching variables, we aimed to achieve standardized differences of 0.25 or less. We aimed to meet these balance criteria when comparing Maryland with the overall benchmark and when comparing Maryland with the high/low spending benchmarks as well. But because our analysis includes a step to regression adjust for differences in characteristics that persist after matching, in some cases, we relaxed these standards so that each Maryland PUMA had a minimum of eight matches.

[54] Specifically, we used multivariate regressions to analyze how strong associations were between spending and the covariates shown in Exhibit F.1 among the sample of all PUMAs in Maryland and, separately, outside of Maryland.

Exhibit F.1. We prioritized variables used in matching into high- and low-priority groups

Variable	Definition	Source
High-priority variables[a]		
HCC score	Mean HCC score for all Medicare FFS beneficiaries in the PUMA	Medicare HCC files
RPP	Relative prices of goods and services calculated at the MSA-level and then merged to each PUMA[b]	Bureau of Economic Analysis
Percentage Black categories	Percentage of Medicare FFS beneficiaries in the PUMA who are Black, in the following categories: 0-5%, 5-15%, 15-30%, and 30% or greater	Medicare enrollment data
Median Household Income categories	Median household income within the PUMA, and then grouped in the following three categories: less than $75,000; $75,000 to $125,000; $125,000 or greater	ACS[c] and Medicare enrollment data
Percentage below federal poverty level, adjusted for cost of living[d]	Mean percentage of residents in the PUMA living below the federal poverty level adjusted for cost of living using the Census supplemental poverty measure	ACS and Census supplemental poverty measure
Rural residence	Percentage of the population living in a rural area in the PUMA	Medicare enrollment data and Census urban and rural classification
Log population density	Logged average population density (persons per square mile) of the PUMA	ACS
Low-priority variables		
Sex	Percentage of all Medicare FFS beneficiaries in the PUMA who are female	Medicare enrollment data
Percentage non-Hispanic White[d]	Percentage of all Medicare FFS beneficiaries in the PUMA who are non-Hispanic White	Medicare enrollment data
Percentage Hispanic	Mean percentage of Hispanic residents in the PUMA	ACS
Age[d]	Mean age of all Medicare FFS beneficiaries in the PUMA	Medicare enrollment data
Percentage disabled or with ESRD[d]	Percentage of Medicare FFS beneficiaries in the PUMA with original reason for entitlement of disability or ESRD	Medicare enrollment data
Number of PCPs per 1,000 Medicare beneficiaries[d]	Total number of PCPs practicing in the PUMA divided by the total number of Medicare beneficiaries (FFS and Medicare Advantage) and multiplied by 1,000	Medicare FFS claims and enrollment data

Note: All variables were measured in 2019

[a] Variables in bold indicate those that we found through empirical analysis to be most strongly associated with spending levels.

[b] Each PUMA merged to a single MSA. For PUMAs corresponding to non-metropolitan areas, we assigned the average non-metropolitan RPP value for the state where the PUMA is located.

[c] Specifically, we used data from the ACS 2019 five-year sample.

[d] Ultimately, we did not use these variables directly in matching but did use them for assessing balance after matching.

ACS = American Community Survey; ESRD = end-stage renal disease; FFS = fee for service; HCC = hierarchical condition category; MSA = metropolitan statistics area; PCP = primary care provider; PUMA = public use microdata area; RPP = regional price parity prospective payment system.

F.3.3. Creating the overall matched benchmark group

To construct the overall benchmark group, we used optimal N:1 matching (with replacement) that minimized global Mahalanobis distances and included calipers. This approach corresponded to 44 separate matching runs (one for each Maryland PUMA) with the same pool of potential matches from across the country used for each problem. Matching each Maryland PUMA separately allowed us to split the matched sets into higher and lower spending groups and retain reasonable balance on the matching variables between Maryland and the high and low spending benchmarks. Although we matched at the individual PUMA level, a primary goal of our approach was to get good balance at the *state level* (that is, between Maryland as a whole and the matched benchmark group and between Maryland and the high and low benchmark groups). We do not estimate how spending would change in individual PUMAs within Maryland if Maryland switched to PPS.

F.3.3a Mahalanobis distances

Mahalanobis distance measures similarity between observations by directly comparing the matching variables in terms of standard deviations, with smaller standard deviations indicating greater similarity. By minimizing Mahalanobis distances of our matching characteristics within each matching problem, we identified sets of PUMAs from across the country that resemble each Maryland PUMA.

F.3.3b Calipers

We used calipers on our matching variables to improve our ability to identify appropriate matches. For each matching problem, if a potential match had a characteristic value that was outside the caliper sizes we imposed (measured in terms of standardized differences), then we dropped that PUMA from consideration. By matching each Maryland PUMA separately, we were able to vary the caliper sizes across each matching problem. In some cases, calipers were loosened or removed to make matching feasible.

F.3.3c Optimal matching

We used optimal N:1 matching (with replacement) to match each PUMA in Maryland to a set of PUMAs outside Maryland. Optimal matching minimizes the total Mahalanobis distance across all matching characteristics between Maryland and matched PUMAs within a set. In each matching problem, we initially required 10 to 20 matches for each set. However, there were 14 Maryland PUMAs for which this requirement was infeasible based on our caliper constraints. For these PUMAs, we lowered the minimum matching requirement to eight matches. We also developed matching weights for assessing balance and use in regression analyses. Maryland PUMAs received a matching weight of 1 and benchmark PUMAs received a weight equal to 1 divided by the number of total benchmark PUMAs in the matched set.

F.3.4. Approach for developing low- and high-spending benchmarks

To develop the low- and high-spending benchmarks, we subdivided the overall benchmarks based on whether total per-capita Medicare FFS spending for the PUMA was below or above the median spending amount measured among all benchmark PUMAs within the matched set. PUMAs that were below (above) the median value were placed in the low (high) group. If a PUMA had the same spending as the median value, then we assigned that PUMA to both the low and high groups. This approach guaranteed an equal number of PUMAs in each group. Matching individual PUMAs allowed us to easily separate matched comparisons into high- and low-spending benchmarks.

F.3.5. Assessing the quality of the matched benchmark group in terms of balance, number of matches, and geographic distribution

In selecting the benchmark groups, we aimed to achieve:

1. Sufficient balance on matching characteristics, particularly those we classified as high priority.
2. A large and geographically diverse set of matches in order to establish a representative national benchmark and to avoid the possibility that idiosyncratic differences that affect spending levels in any one area would drive the results.

Conditions (1) and (2) presented tradeoffs, with achieving more precise balance coming at the expense of a smaller and less geographically dispersed benchmark group. In consultation with CMS, we finalized a benchmark group that best achieved both criteria.

F.3.5a Balance

We achieved good balance between Maryland and the overall benchmark group (Exhibit F.2) as well as between Maryland and the low- and high-spending groups (Exhibit F.3). In the overall benchmark group, all but one high-priority variable (average percentage of population living in a rural area) had a standardized difference of 0.16 or less. Balance for high-priority variables between Maryland and the low- and high-spending groups was slightly worse, but most variables had standardized differences less than 0.25.

Exhibit F.2. We achieved good balance between Maryland and the overall benchmark group

Variables	Standardized difference, pre-matching	Maryland post-matching mean	Benchmark post-matching mean	Standardized difference, post-matching
High priority				
Average HCC score	-0.26	1.14	1.15	-0.16
RPP	0.66	103.25	102.69	0.09
Percentage non-Hispanic Black 0 to 5	-0.80	0.20	0.26	-0.13
Percentage non-Hispanic Black 5 to 15	0.21	0.32	0.31	0.03
Percentage non-Hispanic Black 15 to 30	0.42	0.23	0.22	0.04
Percentage non-Hispanic Black 30 or greater	0.68	0.25	0.22	0.06
Median household income less than $75,000	-0.82	0.19	0.20	-0.02
Median household income $75,000 to $125,000	0.61	0.63	0.65	-0.03
Median household income $125,000 or greater	0.44	0.18	0.15	0.07
Percentage below 100% of the federal poverty level (adjusted for cost of living)	0.02	0.11	0.11	0.12
Log population density	0.45	7.47	7.65	-0.15
Average percentage of the population living in a rural area in the PUMA	-0.36	16.44	11.79	0.29
Low priority				
Average age of FFS patients	0.34	71.67	71.12	0.28
Percentage female	1.26	57.29	55.71	0.85[a]
Percentage non-Hispanic White	-0.72	66.30	69.66	-0.16
Percentage Hispanic	-0.35	0.09	0.11	-0.33
Percentage with original reason for Medicare entitlement: disability, ESRD	-0.36	19.80	21.46	-0.19
Number of PCPs per 1,000 beneficiaries	0.18	5.64	6.54	-0.10

Note: All variables were measured in 2019.

[a] The standardized difference on percentage female is high because the standard deviation for this characteristic is very low (less than 2 percentage points in both the Maryland and benchmark group).

ESRD = end-stage renal disease; FFS = fee for service; HCC = hierarchical condition category; PCP = primary care provider; PUMA = public use microdata area; RPP = regional price parity prospective payment system.

Exhibit F.3. Balance in high- and low-spending benchmark groups

Variables	High-spending benchmark			Low-spending benchmark		
	Maryland post-matching mean	Benchmark post-matching mean	Standardized difference, post-matching	Maryland post-matching mean	Benchmark post-matching mean	Standardized difference, post-matching
High priority						
Average HCC score	1.14	1.16	-0.21	1.14	1.14	-0.10
RPP	103.25	103.36	-0.02	103.25	101.78	0.22
Percentage non-Hispanic Black 0 to 5	0.20	0.28	-0.18	0.20	0.24	-0.09
Percentage non-Hispanic Black 5 to 15	0.32	0.30	0.05	0.32	0.32	0.00
Percentage non-Hispanic Black 15 to 30	0.23	0.22	0.03	0.23	0.20	0.06
Percentage non-Hispanic Black 30 or greater	0.25	0.21	0.10	0.25	0.24	0.02
Median household income less than $75,000	0.19	0.20	-0.02	0.19	0.20	-0.03
Median household income $75,000 to $125,000	0.63	0.64	-0.01	0.63	0.65	-0.03
Median household income $125,000 or greater	0.18	0.16	0.05	0.18	0.15	0.09
Percentage below 100% of the federal poverty level (adjusted for cost of living)	0.11	0.11	0.11	0.11	0.11	0.13
Log population density	7.47	7.68	-0.17	7.47	7.61	-0.12
Average percentage of the population living in a rural area in the PUMA	16.44	11.37	0.32	16.44	12.54	0.23
Low priority						
Average age of FFS patients	71.67	71.30	0.20	71.67	70.96	0.35
Percentage female	57.29	55.98	0.71[a]	57.29	55.39	1.00[a]
Percentage non-Hispanic White	66.30	68.69	-0.10	66.30	70.62	-0.22
Percentage Hispanic	0.09	0.12	-0.41	0.09	0.10	-0.23

Variables	High-spending benchmark			Low-spending benchmark		
	Maryland post-matching mean	Benchmark post-matching mean	Standardized difference, post-matching	Maryland post-matching mean	Benchmark post-matching mean	Standardized difference, post-matching
Percentage with original reason for Medicare entitlement: disability/ESRD	19.80	21.23	-0.17	19.80	21.66	-0.20
Number of primary care providers per 1,000 beneficiaries	5.64	6.15	-0.06	5.64	7.16	-0.17

Note: All variables were measured in 2019.

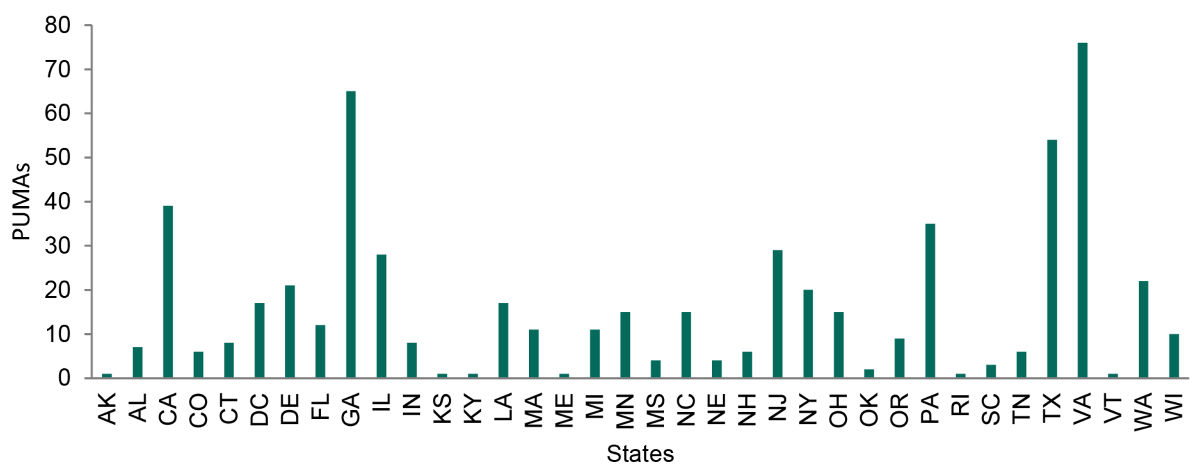
[a] The standardized difference on percentage female is high because the standard deviation for this characteristic is very low (less than 2 percentage points in both the Maryland and benchmark group).

ESRD = end-stage renal disease; FFS = fee for service; HCC = hierarchical condition category; PCP = primary care provider; PUMA = public use microdata area; RPP = regional price parity; PPS = prospective payment system.

F.3.5b Size and geographic spread of the selected benchmark group

Our benchmark group is large and covers much of the country. The overall group consists of 581 total PUMAs, 305 of which are unique PUMAs (some were matched to multiple Maryland PUMAs). On average, each Maryland PUMA was matched to 13 benchmark PUMAs, and all Maryland PUMAs had at least eight matched comparisons. The benchmark group covers 36 states, representing 13% of all PUMAs in the country (Exhibit F.4). Nearly a quarter of the total PUMAs selected are in states with close geographic proximity to Maryland, including Virginia, Pennsylvania, and New Jersey.

Exhibit F.4. The 581 PUMAs that comprise the overall benchmark group are distributed over 36 states, with Virginia contributing more PUMAs than any other state



PUMA = public use microdata area

F.4. Estimating spending and service use effects of switching to PPS

After constructing the benchmark groups by implementing the matching procedure described in the previous section using 2019 data, we estimated spending and service use effects of Maryland switching to PPS using a beneficiary-level data set constructed from Medicare claims files in 2022. Using 2022 data for outcomes is advantageous because it allows us to estimate spending and service use differences between Maryland and the benchmark group that capture several years of the MD TCOC period. We limited the data to beneficiaries who are enrolled in FFS, are observable in at least one month of the year, and reside in Maryland or the PUMAs in the benchmark groups.

Our strategy compared actual mean outcomes in Maryland in 2022 (derived directly from our data) to *predicted* mean outcomes in Maryland under PPS generated using a regression-based approach. We could have estimated spending and service use effects by simply comparing the actual mean outcomes in Maryland and the benchmark groups, but even after matching, there remain some important residual differences between Maryland and the benchmark groups (for example, the percentage of the population living in a rural area as shown in Exhibit F.2). These residual differences could bias our estimates if they are correlated with outcomes. Therefore, we use regressions to correct for these remaining differences on observable characteristics.

To generate predicted mean outcomes, we first estimated the relationship between covariates listed in Exhibit F.5 and outcomes for beneficiaries living in the benchmark areas. (Estimating the relationship between covariates and outcomes for beneficiaries in Maryland, where the all-payer system and hospital global budgets exist, is not relevant). Specifically, we ran a single regression for each outcome and benchmark group (overall, low, high) specified as

$$(1) \qquad y_{ij} = \alpha_i + X_i\beta + X_j\delta + \varepsilon_i$$

where y_i represents the outcome for beneficiary *i* in PUMA *j*, α_i is the intercept, X_i is the set of covariates measured at the beneficiary level, such as HCC scores, in 2022, and X_j are area-level characteristics, such as population density and median household income, measured in 2019.[55] We weighted beneficiary observations using a composite final weight equal to the product of an observability weight and a matching weight. The observability weight, equal to the proportion of 2022 that a beneficiary is eligible for the analytic population, is intended to ensure that beneficiaries who are eligible in 2022 for longer periods contribute more to our estimation than those who are eligible for shorter periods. We used the matching weight to account for the fact that not all matched sets had the same number of non-Maryland PUMAs. The matching weight equals 1 divided by the number of benchmark PUMAs in a matched set. For example, if nine PUMAs were matched to a single Maryland PUMA, the matching weight would equal 1/9.

[55] Most area-level data were not available for 2022 at the time of running the analyses, so we used the same area-level data that we used for the matching procedure measured in 2019.

We then generated the predicted mean outcomes for Maryland beneficiaries after shifting to PPS by adjusting actual outcomes in 2022 by the coefficient estimates generated from equation 1 and aggregating the predicted values to the mean within Maryland, weighting by observability.[56] The projected change in per-capita outcomes in Maryland after switching to PPS is equal to the difference between the predicted mean outcome and actual mean outcome in Maryland in 2022. We calculated estimates of *percentage* differences by dividing the projected per-capita change by the actual mean outcome in Maryland in 2022. For spending outcomes, the projected annual dollar savings are equal to the projected per-capita change multiplied by the weighted number of Maryland beneficiaries in 2022 (N = 721,580).

We did not include corresponding estimates of *statistical significance* with our results for two reasons. First, our sample sizes are large enough that estimates of projected changes in outcomes would all likely be statistically significant. Second, we argue the uncertainty associated with predicting Maryland's *future path after switching to PPS* (because of, for example, differences in unobserved characteristics between Maryland and the benchmarks that affect spending) vastly outweighs capturing the statistical uncertainty of our estimates. This reason motivated our use of the low- and high-spending benchmarks to produce a range of estimates for each outcome.

Exhibit F.5. Beneficiary- and area-level characteristics used to generate predicted outcomes if Maryland shifts to PPS

Variables	Data source
Beneficiary-level characteristics	
Non-Hispanic Black 0/1 flag	Medicare enrollment data
Non-Hispanic White 0/1 flag	Medicare enrollment data
Rural residence 0/1 flag	Medicare enrollment data and Census ZCTA urban and rural classification
HCC score	Medicare HCC files
Age categories Less than 65 65 to 69 70 to 74 75 to 79 80 to 84 85 and older	Medicare enrollment data
Sex (0/1 flag for female)	Medicare enrollment data
Disabled or ESRD 0/1 flag	Medicare enrollment data

[56] When constructing spending outcomes, we first subtracted direct graduate medical expense amounts using data from Medicare cost reports from Maryland claims because these expenses are not part of claim payments for claims outside Maryland.

Variables	Data source
Area-level characteristics	
Population density	ACS
Regional price parity	BEA
Median household income	ACS
Percentage of population below federal poverty level, adjusted for cost of living	ACS and Census supplemental poverty measure
Percentage Hispanic	ACS
Number of PCPs per 1,000 Medicare beneficiaries	Medicare FFS claims and enrollment data

Notes: (1) In this exhibit, we show characteristics (and data sources) used to generate predicted outcomes if Maryland shifted to PPS. Beneficiary-level characteristics were measured in 2022. (2) Area-level characteristics were measured in 2019.

ACS = American Community Survey; BEA = Bureau of Economic Analysis; ESRD = end-stage renal disease; FFS = fee for service; HCC = hierarchical condition category; PCP = primary care provider; ZCTA = zip code tabulation area.

F.5. Supplemental results (comparing results from 2022 with those from 2019)

We assessed whether the results for key outcomes (total spending, inpatient price per stay, and inpatient acute care stays) presented in Chapter 6 were robust to a period before the COVID-19 pandemic by repeating our estimation process but using beneficiaries, covariates, and outcomes in Maryland and the matched comparison PUMAs measured in 2019. We did not make any changes to the matched benchmark groups or the steps we use to estimate projected changes from Maryland shifting to PPS.

Exhibit F.6 shows a comparison of the 2019 and 2022 results. We estimate total spending reductions of more than 10% relative to the Maryland mean per-capita spending in 2019, with projected annual savings totaling more than $1 billion. Similar to our 2022 results, inpatient hospital prices are expected to fall substantially (20%) relative to the projected increase in acute care stays (7%). The differences we observe between 2019 and 2022 results are consistent with findings from the analyses of the Maryland Model over the same period (Chapter 2). For example, hospital prices have continued to increase in Maryland while the MD TCOC Model hasn't further reduced admissions since 2019 relative to the counterfactual. As a result, the price declines from Maryland switching to PPS in 2022 should be larger than they would have been in 2019, which is consistent with results in Exhibit F.6 (showing an estimated 25% decline in price per inpatient stay using 2022 data versus a 20% decline using the 2019 data).

In sum, similar results using data in 2019 and 2022 provides evidence that COVID-19 is not leading to spurious conclusions. Because of the uncertainty of how outcomes will evolve in Maryland and the rest of the country over time, however, comparing results from 2019 and 2022 **should not** be extrapolated to predict what Medicare savings would be from switching Maryland to PPS in future years. For example, we can't simply predict that projected savings of switching to PPS would continue to grow over time. If the Maryland Model succeeds in limiting hospital spending growth more in the future, the projected savings from switching to PPS could remain the same or even decline over time.

Exhibit F.6. The projected change in Medicare FFS spending, inpatient hospital prices, and service use if Maryland switched to Medicare PPS are similar in 2019 versus 2022

Outcome	Percentage change, mean (range) (1)	Projected annual Medicare savings (in $ millions) (2)
2019 beneficiary sample		
Total Part A and B spending	-10.3% (-15.5%, -5.4%)	$1,034 ($1,552, $545)
Price per inpatient stay ($ per stay)	-19.7% (-23.8%, -16.2%)	
Inpatient acute care stays (# per beneficiary)	7.3% (5.4%, 9.6%)	
2022 beneficiary sample		
Total Part A and B spending	-12.6% (-17.6%, -7.8%)	$1,319 ($1,853, $816)
Price per inpatient stay ($ per stay)	-24.7% (-28.3%, -21.7%)	
Inpatient acute care stays (# per beneficiary)	5.7% (4.0%, 7.6%)	

Source: Mathematica's analysis of 2019 and 2022 Medicare claims data.

Notes: (1) In this exhibit, we report the projected change in total (Part A and B) spending, price per inpatient stay, and inpatient acute care stays if Maryland switched to Medicare PPS using a beneficiary sample from 2019 and 2022. (2) Column 1 shows the estimated percentage change relative to actual spending in Maryland during the same year. (3) Column 2 shows the estimated annual Medicare savings (in $ millions).

FFS = fee for service; PPS = prospective payment system.

References

Agency for Healthcare Research and Quality (AHRQ). "Prevention Quality Indicators Overview." n.d.[a] https://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx#techspecs. Accessed September 2023.

AHRQ. "Elixhauser Comorbidity Software, Version 3.7." n.d.[b]. https://hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp. Accessed September 13, 2023.

AHRQ. "Medical Expenditure Panel Survey: Using Appropriate Price Indices for Analyses of Health Care Expenditures or Income Across Multiple Years." n.d.[c]. https://meps.ahrq.gov/about_meps/Price_Index.shtml. Accessed August 18, 2023

AHRQ. "Pediatric Quality Indicator Measures." n.d.[b]. https://qualityindicators.ahrq.gov/measures/pdi_resources. Accessed August 18, 2023.

AHRQ. "Tools Archive for Elixhauser Comorbidity Software Refined for ICD-10-CM." n.d.[c]. https://hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10_archive.jsp. Accessed September 13, 2023.

Billings, John., Geoffrey M. Anderson, and Laurie S. Newman. "Recent Findings on Preventable Hospitalizations." Health Affairs, vol. 15, no. 3, 1996, pp. 239–249. https://doi.org/10.1377/hlthaff.15.3.239

Billings, John, Nina Parikh, and Tod Mijanovich. "Emergency Department Use in New York City: A Substitute for Primary Care?" Issue Brief. New York: Commonwealth Fund and NYU Center for Health and Public Service Research, 2000a. https://wagner.nyu.edu/faculty/billings/nyued-articles.

Billings, John, Nina Parikh, and Tod Mijanovich. "Emergency Department Use: The New York Story." Issue Brief. New York: Commonwealth Fund and NYU Center for Health and Public Service Research, 2000b. https://wagner.nyu.edu/faculty/billings/nyued-articles.

Carmichael, H., A. Moore, L. Steward, and C.G. Velopulos. "Disparities in Emergency Versus Elective Surgery: Comparing Measures of Neighborhood Social Vulnerability." *Journal of Surgical Research,* vol. 256, 2020, pp. 397–403.

Center for Health Disparities Research. "About the Neighborhood Atlas®." n.d. https://www.neighborhoodatlas.medicine.wisc.edu/. Accessed September 6, 2023.

Centers for Disease Control and Prevention (CDC). "CDC Social Vulnerability Index (CDC SVI): A Tool to Identify Socially Vulnerable Communities." n.d. https://www.atsdr.cdc.gov/placeandhealth/svi/fact_sheet/pdf/SVI_FactSheet_v10152020-H.pdf. Accessed July 28, 2022.

CMS. "Maryland Total Cost of Care Model Innovation Center Investment Plan (ICIP). Version 3.0." 2018.

CMS. "MDPCP Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance- Based Incentive Payment, and Comprehensive Primary Care Payment, Model Year 2022." December 2021a.

CMS. "Hospital Compare." Updated December 1, 2021b. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare. Accessed August 2023.

CMS. "HEART Payment Playbook." December 2022a. https://health.maryland.gov/mdpcp/Documents/MDPCP_HEART_Payment_Playbook.pdf.

CMS. "Restructured BETOS Classification System." Updated May 19, 2022b. https://data.cms.gov/provider-summary-by-type-of-service/provider-service-classifications/restructured-betos-classification-system. Accessed August 2022.

CMS. "Acute Inpatient PPS." Updated August 2, 2022c. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS. Accessed August 2022.

CMS. "Case Mix Index." n.d. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Acute-Inpatient-Files-for-Download-Items/CMS022630. Accessed August 2022.

CMS Medicaid and CHIP Business Information Solution (MACBIS). "Maternal and Infant Health (MIH): Identifying Pregnant and Postpartum Beneficiaries in Medicaid and CHIP Administrative Data Reference codes." 2023. https://www.medicaid.gov/sites/default/files/2023-05/mih_reference_codes.pdf. Accessed August 18, 2023.

Chattopadhyay, Ambarish, Christopher H. Hase, and José R. Zubizarreta. "Balancing vs Modeling Approaches to Weighting in Practice." *Statistics in Medicine*, vol. 39, no. 24, October 2020, pp. 3227–3254.

Dafny, L.S. "How Do Hospitals Respond to Price Changes?" *American Economic Review*, vol. 95, no. 5, 2005, pp. 1525–1547.

Daw, J. R., & Hatfield, L. A. (2018). Matching and Regression to the Mean in Difference-in-Differences Analysis. *Health Services Research,* 53(6), 4138–4156. doi: 10.1111/1475-6773.12993.

Erzouki, F. "No Notice of PHE End Gives States Time to Prepare for Medicaid Unwinding." Center on Budget and Policy Priorities, August 16, 2022. https://www.cbpp.org/blog/no-notice-of-phe-end-gives-states-time-to-prepare-for-medicaid-unwinding. Accessed August 17, 2023.

Haft, Howard M., Chad Perman, and Eli Y. Adashi. "The Maryland Primary Care Program—A Blueprint for the Nation?" JAMA, vol. 1, no 10, 2020, e201326. Available at https://jamanetwork.com/journals/jama-health-forum/fullarticle/2772314.

Hannan, E.L., Y. Wu, K. Cozzens, and B. Anderson. "The Neighborhood Atlas Area Deprivation Index For Measuring Socioeconomic Status: An Overemphasis On Home Value." *Health Affairs*, vol. 42, no. 5, 2023. doi: 10.1377/hlthaff.2022.01406.

Harrison E., L. Hula, A. Mims, and M. Alalouf. "Maternal and Infant Health: Identifying Pregnant and Postpartum Beneficiaries Technical Specifications." Centers for Medicare & Medicaid Services, January 2023. https://www.medicaid.gov/sites/default/files/2023-05/mih_techspecs.pdf. Accessed August 18, 2023.

Health Services Cost Review Commission (HSCRC). "Maryland Total Cost of Care Model Statement Agreement." 2018. https://hscrc.maryland.gov/Documents/Modernization/TCOC-State-Agreement-CMMI-FINAL-Signed-07092018.pdf.

HSCRC. "Statewide Integrated Health Improvement Strategy Proposal." Baltimore, MD: HSCRC, December 2020. https://hscrc.maryland.gov/Documents/Modernization/SIHIS%20Proposal%20%20CMMI%20Submission%2012142020.pdf.

HSCRC. "Revised Final Recommendation for the Update Factors for Rate Year 2022." June 2021. https://hscrc.maryland.gov/Documents/Hospitals/gbr-tpr-update/FY-2022/RY22RevisedFinalUFRecommendation_FINAL.pdf. Accessed September 20, 2023.

HSCRC. "EQIP Subgroup January Meeting." January 2022a. https://www.medchi.org/Portals/18/Files/Practice%20Services/EQIP/EQIP%20Subgroup_January_20220128.pdf?ver=2022-01-28-064359-327.

HSCRC. "599th Meeting of the Health Services Cost Review Commission." October 12, 2022b. https://hscrc.maryland.gov/Documents/October%202022%20Public%20Post-Meeting%20Materials%20FINAL.pdf.

HSCRC. "Statewide Integrated Health Improvement Strategy (SIHIS) Annual Report." January 2023. https://hscrc.maryland.gov/Documents/Modernization/Statewide%20Integrated%20Health%20Improvement%20Strategy/SIHIS%202022%20Annual%20Report%20FINAL%20w%20appendices.pdf.

HSCRC. "Readmission Reduction Incentive Program (RRIP)." n.d. https://hscrc.maryland.gov/Pages/init-readm-rip.aspx.

Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS). "Mode & Patient-Mix Adj." Updated August 2022. https://hcahpsonline.org/en/mode--patient-mix-adj/. Accessed August 2022.

IMPAQ Health. "Timely Follow-Up After Acute Exacerbations of Chronic Conditions." 2018. https://impaqint.com/measure-information-timely-follow-after-acute-exacerbations-chronic-conditions. Accessed September 2021.

Johnston, Kenton J., Lindsay Allen, Taylor A. Melanson, and Stephen R. Pitts. "A 'Patch' to the NYU Emergency Department Visit Algorithm." *Health Services Research,* vol. 52, no. 4, 2017, pp. 1264–1276.

Maryland Health Services Cost Review Commission (HSCRC). "Quality Initiatives." n.d. https://hscrc.maryland.gov/Pages/quality.aspx. Accessed September 6, 2023.

Mukamel, D.B., H. Ladd, Y. Li, H. Temkin-Greener, and Q. Ngo-Metzger. "Have Racial Disparities in Ambulatory Care Sensitive Admissions Abated Over Time?" *Medical Care,* vol. 53, no. 11, 2015, pp. 931–939.

Peikes, Deborah, Kaylyn Swankoski, Ann O'Malley, Lori Timmins, Dana Petersen, Kristin Geonnotti, Ha Tu, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+). Third Annual Report." Princeton, NJ: Mathematica, 2021. Available at https://innovation.cms.gov/data-and-reports/2021/cpc-plus-third-anual-eval-report.

Peterson, C., and D. Schumacher. "How Maryland's Total Cost Of Care Model Has Helped Hospitals Manage The COVID-19 Stress Test." Health Affairs Blog, October 7, 2020. Available at https://www.healthaffairs.org/do/10.1377/forefront.20201005.677034/full/.

Perman, C., E. Adashi, E. Gruber, and H. Haft. "Improving COVID-19 Outcomes for Medicare Beneficiaries: A Public Health-Supported Advanced Primary Care Paradigm." New York, NY: The Milbank Memorial Fund, September 9, 2021. Available at https://www.milbank.org/publications/improving-covid-19-outcomes-for-medicare-beneficiaries-a public-health-supported-advanced-primary-care-paradigm/.

Pines, J.M., S. Vats, M.S. Zocchi, and C. Black. "Maryland's Experiment with Capitated Payments for Rural Hospitals: Large Reductions in Hospital-Based Care." *Health Affairs*, vol. 38, no. 4, 2019, pp. 594–603

Rotter, Jason, Keri Calkins, Kate Stewart, Isabel Platt, Rachel Machta, Keith Kranker, Nancy McCall, and Greg Peterson. "Evaluation of the Maryland Total Cost of Care Model: Quantitative-Only Report for the Model's First Three Years (2019 to 2021)." Mathematica, 2022. https://www.cms.gov/priorities/innovation/data-and-reports/2022/md-tcoc-qor2.

Stuart, Elizabeth A. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science*, vol. 25, no. 1, February 2010, pp. 1–21.

Wang, Yixin, and José R. Zubizarreta. "Minimal Dispersion Approximately Balancing Weights: Asymptotic Properties and Practical Considerations." *Biometrika*, vol. 107, no. 1, October 2019.

Wikle, S., and J. Wagner. "Unwinding the Medicaid Continuous Coverage Requirement: Frequently Asked Questions." Center on Budget and Policy Priorities, March 9, 2022. https://www.cbpp.org/research/health/unwinding-the-medicaid-continuous-coverage-requirement. Accessed August 17, 2023.

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation. "2020 Readmission Measures Updates and Specifications Reports and Supplemental File." 2020. https://qualitynet.cms.gov/inpatient/measures/readmission/resources#tab3. Accessed August 2023.

Zubizarreta, José R. "Stable Weights that Balance Covariates for Estimation with Incomplete Outcome Data." *Journal of the American Statistical Association*, vol. 110, no. 511, 2015, pp. 910–922.

Mathematica Inc.

Our employee-owners work nationwide and around the world.

Find us at **mathematica.org** and **edi-global.com**.